GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 23, 2005, 23:08:06 ; Search time 763.417 Seconds
(without alignments)
1396.372 Million cell updates/sec

Title: US-10-848-737-1
Perfect score: 22
Sequence: 1 gugaacucacucgugagcuctt 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	90.9	1706	14	AY322205S1	AY322205 SARS coro
2	20	90.9	14011	14	AY534758S1	AY534758 SARS coro
3	20	90.9	26333	14	AY286320	AY286320 SARS coro
4	20	90.9	29013	14	AY463060	AY463060 SARS coro
5	20	90.9	29350	14	AY394999	AY394999 SARS coro
6	20	90.9	29350	14	AY395000	AY395000 SARS coro
7	20	90.9	29350	14	AY395001	AY395001 SARS coro
8	20	90.9	29350	14	AY395002	AY395002 SARS coro
9	20	90.9	29433	14	AY394977	AY394977 SARS coro
10	20	90.9	29530	14	AY394985	AY394985 SARS coro
11	20	90.9	29573	14	AY338174	AY338174 SARS coro
12	20	90.9	29573	14	AY338175	AY338175 SARS coro
13	20	90.9	29573	14	AY348314	AY348314 SARS coro
14	20	90.9	29577	14	AY559094	AY559094 SARS coro
15	20	90.9	29592	14	AY463059	AY463059 SARS coro
16	20	90.9	29620	14	AY395004	AY395004 SARS coro
17	20	90.9	29640	14	AY394978	AY394978 SARS coro
18	20	90.9	29645	14	AY394979	AY394979 SARS coro
19	20	90.9	29646	14	AY394982	AY394982 SARS coro
20	20	90.9	29647	14	AY395003	AY395003 SARS coro
21	20	90.9	29661	14	AY559086	AY559086 SARS coro
22	20	90.9	29665	14	AY394988	AY394988 SARS coro
23	20	90.9	29670	14	AY559082	AY559082 SARS coro
24	20	90.9	29683	14	AY394996	AY394996 SARS coro
25	20	90.9	29683	14	AY394997	AY394997 SARS coro
26	20	90.9	29699	14	AY394983	AY394983 SARS coro
27	20	90.9	29705	14	AY283795	AY283795 SARS coro
28	20	90.9	29705	14	AY394980	AY394980 SARS coro
29	20	90.9	29706	14	AY283797	AY283797 SARS coro
30	20	90.9	29709	14	AY394987	AY394987 SARS coro
31	20	90.9	29710	14	AY559091	AY559091 SARS coro
32	20	90.9	29711	14	AY283794	AY283794 SARS coro
33	20	90.9	29711	14	AY283796	AY283796 SARS coro
34	20	90.9	29711	14	AY283798	AY283798 SARS coro
35	20	90.9	29711	14	AY427439	AY427439 SARS coro
36	20	90.9	29712	14	AY559093	AY559093 SARS coro
37	20	90.9	29713	14	AY559085	AY559085 SARS coro
38	20	90.9	29713	14	AY559092	AY559092 SARS coro
39	20	90.9	29714	14	AY559088	AY559088 SARS coro
40	20	90.9	29715	14	AY297028	AY297028 SARS coro
41	20	90.9	29715	14	AY461660	AY461660 SARS coro
42	20	90.9	29715	14	AY559097	AY559097 SARS coro
43	20	90.9	29716	14	AY559081	AY559081 SARS coro
44	20	90.9	29716	14	AY559087	AY559087 SARS coro
45	20	90.9	29716	14	AY595412	AY595412 SARS coro

ALIGNMENTS

RESULT 1
AY322205S1
LOCUS AY322205S1 1706 bp RNA linear VRL 21-JUL-2003
DEFINITION SARS coronavirus Shanhgai LY orf1ab polyprotein and orf1a polyprotein genes, partial cds.
ACCESSION AY322205
VERSION AY322205.1 GI:32454339
KEYWORDS .
SEGMENT 1 of 4
SOURCE SARS coronavirus Shanhgai LY
ORGANISM SARS coronavirus Shanhgai LY
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales; Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 1706)
AUTHORS Yuan,Z., Zhang,X., Hu,Y., Lan,S., Wang,H., Zhou,Z. and Wen,Y.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2003) Molecular Virology, Shanghai Medical College of Fudan University, 138 Yi Xue Yuan Road, Shanghai 200032, P.R. China
FEATURES Location/Qualifiers
source 1. .1706
/organism="SARS coronavirus Shanhgai LY"
/mol_type="genomic RNA"
/db_xref="taxon:235173"
/country="China: Shanghai"
CDS 186. .>1706
/codon_start=1
/product="orf1a polyprotein"
/protein_id="AAP82966.1"
/db_xref="GI:32454344"
/translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEDYEQN
WNTKHGSGALRELTRELNGGVVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
LDYIESKRGVYCCRDHEHEIAWFTERSDKSCEHQTPFEIKSAKKFDTFKGECPKFVFP
LNSKVKVIQPRVEKKKTEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTC
DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSN
IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
LEILSRERVNINIVGDFHLNEEVAIILASFSASTSAFIDTIKSLDYKSFKTIVESCG"
CDS 186. .>1706
/note="translated by ribosome slippage event"
/codon_start=1

```
                     /product="orf1ab polyprotein"
                     /protein_id="AAP82965.1"
                     /db_xref="GI:32454343"
                     /translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
                     HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
                     TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEDYEQN
                     WNTKHGSGALRELTRELNGGVVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
                     LDYIESKRGVYCCRDHEHEIAWFTERSDKSCEHQTPFEIKSAKKFDTFKGECPKFVFP
                     LNSKVKVIQPRVEKKKTEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTC
                     DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSN
                     IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
                     LEILSRERVNINIVGDFHLNEEVAIILASFSASTSAFIDTIKSLDYKSFKTIVESCG"
ORIGIN

  Query Match             90.9%;  Score 20;  DB 14;  Length 1706;
  Best Local Similarity   75.0%;  Pred. No. 5.7;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        697 GTGAACTCACTCGTGAGCTC 716


RESULT 2
AY534758S1
LOCUS       AY534758S1              14011 bp    RNA     linear   VRL 17-MAR-2004
DEFINITION  SARS coronavirus Sin0409, partial sequence.
ACCESSION   AY534758
VERSION     AY534758.1  GI:45384965
KEYWORDS    .
SEGMENT     1 of 4
SOURCE      SARS coronavirus Sin0409
  ORGANISM  SARS coronavirus Sin0409
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 14011)
  AUTHORS   Lim,P.L., Kurup,A., Gopalakrishna,G., Chan,K.P., Wong,C.W.,
            Ng,L.C., Se-Thoe,S.Y., Oon,L., Bai,X., Stanton,L.W., Ruan,Y.,
            Miller,L.D., Vega,V.B., James,L., Ooi,P.L., Kai,C.S., Olsen,S.J.,
            Ang,B. and Leo,Y.S.
  TITLE     Laboratory-acquired severe acute respiratory syndrome (SARS) -
            Singapore 2003
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 14011)
  AUTHORS   Wei,C.L., Lee,C., Lin,S., Thoreau,H., Vega,V.B., Stanton,L.W. and
            Ruan,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JAN-2004) Genome Institute of Singapore, 60, Biopolis
            Street, 02-01, Genome, Singapore 138672, Singapore
FEATURES             Location/Qualifiers
     source          1. .14011
                     /organism="SARS coronavirus Sin0409"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:266147"
                     /country="Singapore"
ORIGIN

  Query Match             90.9%;  Score 20;  DB 14;  Length 14011;
  Best Local Similarity   75.0%;  Pred. No. 6.2;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        231 GTGAACTCACTCGTGAGCTC 250


RESULT 3
AY286320
LOCUS       AY286320                26333 bp    RNA     linear   VRL 09-FEB-2004
DEFINITION  SARS coronavirus ZJ01, partial genome.
ACCESSION   AY286320
VERSION     AY286320.4  GI:39980888
KEYWORDS    .
SOURCE      SARS coronavirus ZJ01
  ORGANISM  SARS coronavirus ZJ01
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 26333)
  AUTHORS   Li,L., Wang,Z., Lu,Y., Bao,Q., Chen,S., Wu,N., Cheng,S., Weng,J.,
            Zhang,Y., Yan,J., Mei,L., Wang,X., Zhu,H., Yu,Y., Zhang,M., Li,M.,
            Yao,J., Lu,Q., Yao,P., Bo,X., Wo,J., Wang,S. and Hu,S.
  TITLE     Severe acute respiratory syndrome-associated coronavirus genotype
            and its characterization
  JOURNAL   Chin. Med. J. 116 (9), 1288-1292 (2003)
  MEDLINE   22889812
   PUBMED   14527350
REFERENCE   2  (bases 1 to 26333)
  AUTHORS   Wang,Z.G., Li,L.J., Luo,Y., Zhang,J.Y., Wang,M.Y., Cheng,S.Y.,
            Zhang,Y.J., Wang,X.M., Lu,Y.Y., Wu,N.P., Mei,L.L. and Wang,Z.X.
  TITLE     Molecular biological analysis of genotyping and phylogeny of severe
            acute respiratory syndrome associated coronavirus
  JOURNAL   Chin. Med. J. 117 (1), 42-48 (2004)
   PUBMED   14733771
REFERENCE   3  (bases 1 to 26333)
  AUTHORS   Wang,Z., Cheng,S. and Zhang,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-APR-2003) Department of Microbiology, Zhejiang CDC,
            17 Laozhedazhi Rd., Hangzhou, Zhejiang 310009, China
REFERENCE   4  (bases 1 to 26333)
  AUTHORS   Wang,Z., Cheng,S. and Zhang,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-2003) Department of Microbiology, Zhejiang CDC,
            17 Laozhedazhi Rd., Hangzhou, Zhejiang 310009, China
  REMARK    Sequence update by submitter
REFERENCE   5  (bases 1 to 26333)
  AUTHORS   Wang,Z., Cheng,S. and Zhang,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) Department of Microbiology, Zhejiang CDC,
            17 Laozhedazhi Rd., Hangzhou, Zhejiang 310009, China
  REMARK    Sequence update by submitter
REFERENCE   6  (bases 1 to 26333)
  AUTHORS   Wang,Z., Cheng,S. and Zhang,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-DEC-2003) Department of Microbiology, Zhejiang CDC,
            17 Laozhedazhi Rd., Hangzhou, Zhejiang 310009, China
  REMARK    Sequence update by submitter
COMMENT     On Dec 17, 2003 this sequence version replaced gi:38304880.
FEATURES             Location/Qualifiers
     source          1. .26333
                     /organism="SARS coronavirus ZJ01"
                     /mol_type="genomic RNA"
                     /isolate="ZJ01b"
                     /isolation_source="cultured virus in Vero cell inoculated
                     with throat swab from patient with severe acute
                     respiratory syndrome (SARS)"
                     /db_xref="taxon:230471"
                     /country="China: Hangzhou"
                     /note="isolated April 21, 2003"
     5'UTR           1. .248
     misc_feature    51. .56
                     /note="transcription regulatory sequence leader; TRS"
     CDS             join(249. .13382,13382. .21469)
                     /note="ribosomal slippage"
                     /codon_start=1
                     /product="orf1ab polyprotein"
                     /protein_id="AAP49011.4"
                     /db_xref="GI:38324306"
                     /translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
                     HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
                     TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEDYEQN
                     WNTKHGSGALRELTRELNGGAVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
                     LDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPFEIKSAKKFDTFKGECPKFVFP
                     LNSKVKVIQPRVEKKKTEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTC
                     DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSN
                     IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
                     LEILSRERVNINIVGDFHLNEEVAIILASFSASTSAFIDTIKSLDYKSFKTIVESCGN
```

```
YKVTKGKPVKGAWNIGQQRSVLTPLCGFPSQAAGVIRSIFARTLDAANHSIPDLQRAA
VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVQQTSQWLSNLLGTTVEKL
RPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKGQIQVASDNIKDCVKCFIDVV
NKALEMCIDQVTIAGAKLRSLNLGEVFIAQSKGLYRQCIRGKEQLQLLMPLKAPKEVT
FLEGDSHDTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTPVCVNGLMLLEIKDKEQ
YCALSPGLLATNNVFRLKGGAPIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNE
KCSVYTVESGTEVTEFACVVAEAVVKTLQPVSDLLTNMGIDLDEWSVATFYLFDDAGE
ENFSSRMYCSFYPPDEEEEDDAECEEEEIDETCEHEYGTEDDYQGLPLEFGASAETVR
VEEEEEEDWLDDTTEQSEIEPEPEPTPEEPVNQFTGYLKLTDNVAIKCVDIVKEAQSA
NPMVIVNAANIHLKHGGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL
AKKCLHVVGPNLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQT
VRTQVYIAVNDKALYEQVVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPVDV
KPKIKACIDEVTTTLEETKFLTNKLLLFADINGKLYHDSQNMLRGEDMSFLEKDAPYM
VGDVITSGDITCVVIPSKKAGGTTEMLSRALKKVPVDEYITTYPGQGCAGYTLEEAKT
ALKKCKSAFYVLPSEAPNAKEEILGTVSWNLREMLAHAEETRKLMPICMDVRAIMATI
QRKYKGIKIQEGIVDYGVRFFFYTSKEPVASIITKLNSLNEPLVTMPIGYVTHGFNLE
EAARCMRSLKAPAVVSVSSPDAVTTYNGYLTSSSKTSEEHFVETVSLAGSYRDWSYSG
QRTELGVEFLKRGDKIVYHTLESPVEFHLDGEVLSLDKLKSLLSLREVKTIKVFTTVD
NTNLHTQLVDMSMTYGQQFGPTYLDGADVTKIKPHVNHEGKTFFVLPSDDTLRSEAFE
YYHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVLLALQQLEVKFN
APALQEAYYRARAGDAANFCALILAYSNKTVGELGDVRETMTHLLQHANLESAKRVLN
VVCKHCGQKTTTLTGVEAVMYMGTLSYDNLKTGVSIPCVCGRDATQYLVQQESSFVMM
SAPPAEYKLQQGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTD
VFYKETSYTTTIKPVSYKLDGVTYTEIEPKLDGYYKKDNAYYTEQPIDLVPTQPLPNA
SFDNFKLTCSNTKFADDLNQMTGFTKPASRELSVTFFPDLNGDVVAIDYRHYSASFKK
GAKLLHKPIVWHINQATTKTTFKPNTWCLRCLWSTKPVDTSNSFEVLAVEDTQGMDNL
ACESQQPTSEEVVENPTIQKEVIECDVKTTEVVGNVILKPSDEGVKVTQELGHEDLMA
AYVENTSITIKKPNELSLALGLKTIATHGIAAINSVPWSKILAYVKPFLGQAAITTSN
CAKRLAQRVFNNYMPYVFTLLFQLCTFTKSTNSRIRASLPTTIAKNSVKSVAKLCLDA
GINYVKSPKFSKLFTIAMWLLLLSICLGSLICVTAAFGVLLSNFGAPSYCNGVRELYL
NSSNVTTMDFCEGSFPCSICLSGLDSLDSYPALETIQVTISSYKLDLTILGLAAEWVL
AYMLFTKFFYLLGLSAIMQVFFGYFASHFISNSWLMWFIISIVQMAPVSAMVRMYIFF
ASFYYIWKSYVHIMDGCTSSTCMMCYKRNRATRVECTTIVNGMKRSFYVYANGGRGFC
KTHNWNCLNCDTFCTGSTFISDEVARDLSLQFKRPINPTDQSSYIVDSVAVKNGALHL
YFDKAGQKTYERHPLSHFVNLDNLRANNTKGSLPINVIVFDGKSKCDESASKSASVYY
SQLMCQPILLLDQALVSDVGDSTEVSVKMFDAYVDTFSATFSVPMEKLKALVATAHSE
LAKGVALDGVLSTFVSAARQGVVDTDVDTKDVIECLKLSHHSDLEVTGDSCNNFMLTY
NKVENMTPRDLGACIDCNARHINAQVAKSHNVSLIWNVKDYMSLSEQLRKQIRSAAKK
NNIPFRLTCATTRQVVNVITTKISLKGGKIVSTCFKLMLKATLLCVLAALVCYIVMPV
HTLSIHDGYTNEIIGYKAIQDGVTRDIISTDDCFANKHAGFDAWFSQRGGSYKNDKSC
PVVAAIITREIGFIVPGLPGTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEYSDFA
TSACVLAAECTIFKDAMGKPVPYCYDTNLLEGSISYSELRPDTRYVLMDGSIIQFPNT
YLEGSVRVVTTFDAEYCRHGTCERSEVGICLSTSGRWVLNNEHYRALSGVFCGVDAMN
LIANIFTPLVQPVGALDVSASVVAGGIIAILVTCAAYYFMKFRRVFGEYNHVVAANAL
LFLMSFTILCLVPAYSFLPGVYSVFYLYLTFYFTNDVSFLAHLQWFAMFSPIVPFWIT
AIYVFCISLKHCHWFFNNYLRKRVMFNGVTFSTFEEAALCTFLLNKEMYLKLRSETLL
PLTQYNRYLALYNKYKYFSGALDTTSYREAACCHLAKALNDFSNSGADVLYQPPQTSI
TSAVLQSGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDTVYCPRHVICTAEDMLNP
NYEDLLIRKSNHSFLVQAGNVQLRVIGHSMQNCLLRLKVDTSNPKTPKYKFVRIQPGQ
TFSVLACYNGSPSGVYQCAMRPNHTIKGSFLNGSCGSVGFNIDYDCVSFCYMHHMELP
TGVHAGTDLEGKFYGPFVDRQTAQAAGTDTTITLNVLAWLYAAVINGDRWFLNRFTTT
LNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCAALKELLQNGMNGRTILGS
TILEDEFTPFDVVRQCSGVTFQGKFKKIVKGTHHWMLLTFLTSLLILVQSTQWSLFFF
VYENAFLPFTLGIMAIAACAMLLVKHKHAFLCLFLLPSLATVAYFNMVYMPASWVMRI
MTWLELADTSLSGYRLKDCVMYASALVLLILMTARTVYDDAARRVWTLMNVITLVYKV
YYGNALDQAISMWALVISVTSNYSGVVTTIMFLARAIVFVCVEYYPLLFITGNTLQCI
MLVYCFLGYCCCCYFGLFCLLNRYFRLTLGVYDYLVSTQEFRYMNSQGLLPPKSSIDA
FKLNIKLLGIGGKPCIKVATVQSKMSDVKCTSVVLLSVLQQLRVESSSKLWAQCVQLH
NDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINRLCEEMLDNRATLQAIASEFSSLPS
YAAYATAQEAYEQAVANGDSEVVLKKLKKSLNVAKSEFDRDAAMQRKLEKMADQAMTQ
MYKQARSEDKRAKVTSAMQTMLFTMLRKLDNDALNNIINNARDGCVPLNIIPLTTAAK
LMVVVPDYGTYKNTCDGNTFTYASALWEIQQVVDADSKHVQLSEINMDNSPNLAWPLI
VTALRANSAVKLQNNELSPVALRQMSCAAGTTQTACTDDNALAYYNNSKGGRFVLALL
SDHQDLKWARFPKSDGTGTIYTELEPPCRFVTDTPKGPKVKYLYFIKGLNNLNRGMVL
GSLAATVRLQAGNATEVPAHSTVLSFCAFAVDPAKAYKDYLASGGQPITNCVKMLCTH
TGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHPNPKGFCDLKGKYVQIPTTCAND
PVGFTLRNTVCTVCGMWKGYGCSCDQLREPLMQSADASTFLNRVCGVSAARLTPCGTG
TSTDVVYRAFDIYNEKVAGFAKFLKTNCCRFQEKDEEGNLLDSYFVVKRHTMSNYQHE
ETIYNLVKDCPAVAVHDFFKFRVDGDMVPHISRQRLTKYTMADLVYALRHFDEGNCDT
LKEILVTYNCCDDDYFNKKDWYDFVENPDILRVYANLGERVRQSLLKTVQFCDAMRDA
GIVGVLTLDNQDLNGNWYDFGDFEQVAPGCGVPIVDSYYSLLMPILTLTRALAAESHM
DADLAKPLIKWDLLKYDFTEERLCLFDRYFKYWDQTYHPNCINCLDDRCILHCANFNV
LFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVNLHSSRLSFKELLV
```

```
                     YAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFYDFAVSKGFFKE
                     GSSVELKHFFFAQDGNAAISDYDYYRYNLPTMCDIRQLLFVVEVVDKYFDCYDGGCIN
                     ANQVIVNNLDKSAGFPFNKWGKARLYYDSMSYEDQDALFAYTKRNVIPTITLMNLKYA
                     ISAKNRARTVAGVSICSTMTNRQFHQKLLKSIAATRGATVVIGTSKFYGGWHNMLKTV
                     YSDVETPHLMGWDYPKCDRAMPNMLRIMASLVLARKHNTCCNLSHRFYRLANECAQVL
                     SEMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQAVTANVNALLSTDGNKIADKYVR
                     NLQHRLYECLYRNRDVDHEFVDEFYAYLRKHFSMMILSDDAVVCYNSNYAAQGLVASI
                     KNFKAVLYYQNNVFMSEAKCWTETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRI
                     LGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADVFHLYLQYIRKLHDEL
                     TGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTVLQAVGACVLCNSQTSLRCGACIR
                     RPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCDVTDVTQLYLGGMSYYCKSHKPPI
                     SFPLCANGQVFGLYKNTCVGSDNVTDFNAIATCDWTNAGDYILANTCTERLKLFAAET
                     LKATEETFKLSYGIATVREVLSDRELHLSWEVGKPRPPLNRNYVFTGYRVTKNSKVQI
                     GEYTFEKGDYGDAVVYRGTTTYKLNVGDYFVLTSHTVMPLSAPTLVPQEHYVRITGLY
                     PTLNISDEFSSNVANYQKVGMQKYSTLQGPPGTGKSHFAIGLALYYPSARIVYTACSH
                     AAVDALCEKALKYLPIDKCSRIIPARARVECFDKFKVNSTLEQYVFCTVNALPETTAD
                     IVVFDEISMATNYDLSVVNARLRAKHYVYIGDPAQLPAPRTLLTKGTLEPEYFNSVCR
                     LMKTIGPDMFLGTCRCCPAEIVDTVSALVYDNKLKAHKDKSAQCFKMFYKGVITHDVS
                     SAINRPQIGVVREFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQTVDSSQGSEYDY
                     VIFTQTTETAHSCNVNRFNVAITRAKIGILCIMSDRDLYDKLQFTSLEIPRRNVATLQ
                     AENVTGLFKDCSKIITGLHPTQAPTHLSVDIKFKTEGLCVDIPGIPKDMTYRRLISMM
                     GFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEGCHATRDAVGTNLPLQLGFSTGVN
                     LVAVPTGYVDTENNTEFTRVNAKPPPGDQFKHLIPLMYKGLPWNVVRIKIVQMLSDTL
                     KGLSDRVVFVLWAHGFELTSMKYFVKIGPERTCCLCDKRATCFSTSSDTYACWNHSVG
                     FDYVYNPFMIDVQQWGFTGNLQSNHDQHCQVHGNAHVASCDAIMTRCLAVHECFVKRV
                     DWSVEYPIIGDELRVNSACRKVQHMVVKSALLADKFPVLHDIGNPKAIKCVPQAEVEW
                     KFYDAQPCSDKAYKIEELFYSYATHHDKFTDGVCLFWNCNVDRYPANAIVCRFDTRVL
                     SNLNLPGCDGGSLYVNKHAFHTPAFDKSAFTNLKQLPFFYYSDSPCESHGKQVVSDID
                     YVPLKSATCITRCNLGGAVCRHHANEYRQYLDAYNMMISAGFSLWIYKQFDTYNLWNT
                     FTRLQSLENVAYNVVNKGHFDGHAGEAPVSIINNAVYTKVDGIDVEIFENKTTLPVNV
                     AFELWAKRNIKPVPEIKILNNLGVDIAANTVIWDYKREAPAHVSTIGVCTMTDIAKKP
                     TESACSSLTVLFDGRVEGQVDLFRNARNGVLITEGSVKGLTPSKGPAQASVNGVTLIG
                     ESVKTQFNYFKKVDGIIQQLPETYFTQSRDLEDFKPRSQMETDFLELAMDEFIQRYKL
                     EGYAFEHIVYGDFSHGQLGGLHLMIGLAKRSQDSPLKLEDFIPMDSTVKNYFITDAQT
                     GSSKCVCSVIDLLLDDFVEIIKSQDLSVISKVVKVTIDYAEISFMLWCKDGHVETFYP
                     KLQASQAWQPGVAMPNLYKMQRMLLEKCDLQNYGENAVIPKGIMMNVAKYTQLCQYLN
                     TLTLAVPYNMRVIHFGAGSDKGVAPGTAVLRQWLPTGTLLVDSDLNDFVSDADSTLIG
                     DCATVHTANKWDLIISDMYDPRTKHVTKENDSKEGFFTYLCGFIKQKLALGGSIAVKI
                     TEHSWNADLYKLMGHFSWWTAFVTNVNASSSEAFLIGANYLGKPKEQIDGYTMHANYI
                     FWRNTNPIQLSSYSLFDMSKFPLKLRGTAVMSLKENQINDMIYSLLEKGRLIIRENNR
                     VVVSSDILVNN"
     CDS             249. .13397
                     /codon_start=1
                     /product="orf1a polyprotein"
                     /protein_id="AAR16181.1"
                     /db_xref="GI:38304882"
                     /translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
                     HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
                     TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEDYEQN
```

```
  Query Match          90.9%;  Score 20;  DB 14;  Length 26333;
  Best Local Similarity  75.0%;  Pred. No. 6.4;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        760 GTGAACTCACTCGTGAGCTC 779


RESULT 4
AY463060
LOCUS       AY463060               29013 bp    RNA     linear   VRL 05-JAN-2004
DEFINITION  SARS coronavirus ShanghaiQXC2, complete genome.
ACCESSION   AY463060
VERSION     AY463060.1  GI:40457448
KEYWORDS    .
SOURCE      SARS coronavirus ShanghaiQXC2
  ORGANISM  SARS coronavirus ShanghaiQXC2
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29013)
  AUTHORS   Yuan,Z., Zhang,X., Hu,Y., Lan,S., Zhou,Z., Wang,H. and Wen,Y.
  TITLE     Analysis of SARS coronavirus genome in Shanghai isolates
```

```
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 29013)
  AUTHORS   Yuan,Z., Zhang,X., Hu,Y., Lan,S., Zhou,Z., Wang,H. and Wen,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-2003) Key Lab of Medical Molecular Virology,
            Shanghai Medical College, Fudan University, 138 Yi Xue Yuan Road,
            Shanghai 200032, P.R. China
FEATURES             Location/Qualifiers
     source          1. .29013
                     /organism="SARS coronavirus ShanghaiQXC2"
                     /virion
                     /mol_type="genomic RNA"
                     /strain="ShanghaiQXC2"
                     /db_xref="taxon:258508"
     CDS             join(186. .12734,12734. .20827)
                     /note="ribosomal slippage"
                     /codon_start=1
                     /product="orflab"
                     /protein_id="AAR86774.1"
                     /db_xref="GI:40457449"
                     /translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
                     HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
                     TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEDYEQN
                     WNTKHGSGALRELTRELNGGAVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
                     LDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPFEIKSAKKFDTFKGECPKFVFP
                     LNSKVKVIQPRVEKKKTEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTC
                     DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSN
                     IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
                     LEILSRERVNINIVGDFHLNEEVAIILASFSASTSAFIDTIKSLDYKSFXTIVESCGN
                     YKVTKGKPVKGAWNIGQQRSVLTPLCGFPSQAAGVIRSIFARTLDAANHSIPDLQRAA
                     VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVQQTSQWLSNLLGTTVEKL
                     RPIFEWIEAKLSAGVEFLKDAWEILKPLITGVFDIVKGQIQVASDNIKDCVKCFIDVV
                     NKALEMCIDQVTIAGAKLRSLNLGEVFIAQSKGLYRQCIRGKEQLQLLMPLKAPKEVT
                     FLEGDSHDTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTPVCVNGLMLLEIKDKEQ
                     YCALSPGLLATNNVFRLKGGAPIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNE
                     KCSVYTVESGTEVTEFACVVAEAVVKTLQPVSDLLTNMGIDLDEWSVATFYLFDDAGE
                     ENFSSRMYCSFYPPDEEEEDDAECEEEEIDETCEHEYGTEDDYQGLPLEFGASAETVR
                     VEEEEEEDWLDDTTEQSEIEPEPEPTPEEPVNQFTGYLKLTDNVAIKCVDTVKEAQSA
                     NPMVIVNAANIHLKHGGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL
                     AKKCLHVVGPNLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQT
                     VRTQVYIAVNDKALYEQVVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPVDV
                     KPKIKACIDEVTTTLEETKFLTNKLLLFADINGKLYHDSQNMLRGEDMSFLEEDAPYM
                     VGDVITSGDITCVVIPSKKAGGTTEMLSRALKKVPVDEYITTYPGQGCAGYTLEEAKT
                     ALKKCKSAFYVLPSEAPNAKEEILGTVSWNLREMLAHAEEARKLMPICMDVRAIMATI
                     QRKYKGVKIQEGIVDYGVRFFFYTSKEPVASIITKLNSLNEPLVTMPIGYVTHGFNLE
                     EAARCMRSLKAPAVVSVSSPDAVTTYNGYLTSSSKTSEEHFVETVSLAGSYRDWSYSG
                     QRTELGVEFLKRGDKIVYHTLESPVEFHLDGEVLSLDKLKSLLSLREVKTIKVFTTVD
                     NTNLHTQLVDMSMTYGQQFGPTYLDGADVTKIKPHVNHEGKTFFVLPSDDTLRSEAFE
                     YYHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVLLALQQLEVKFN
                     APALQEAYYRARAGDAANFCALILAYSNKTVGELGDVRETMTHLLQHANLESAKRVLN
                     VVCKHCGQKTTTLTGVEAVMYMGTLSYDNLKTGVSIPCVCGRDATQYLVQQESSFVMM
                     SAPPAGYKLQQGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTD
                     VFYKETSYTTTIKPVSYKLDGVTYTEIEPKLDGYYKKDNAYVENTSITIKKPNELSLA
                     LGLKTIATHGIAAINSVPWSKILAYVKPFLGQAAITTSNCAKRLAQRVFNNYMPYVFT
                     LLFQLCTFTKSTNSRIRASLPTTIAKNSVKSVAKLCLDAGINYVKSPKFSKLFTIAMW
                     LLLLSICLGSLICVTAAFGVLLSNFGAPSYCNGVRELYLNSSNVTTMDFCEGSFPCSI
                     CLSGLDSLDSYPALETIQVTISSYKLDLTILGLAAEWVLAYMLFTKFFYLLGLSAIMQ
                     AFFGYFASHFISNSWLMWFIISIVQMAPVSAMVRMYIFFASFYYIWKSYVHIMDGCTS
                     STCMMCYKCNHATRVECTTIVNGMKRSFYVYANGGRGFCKTHNWNCLNCDTFCTGSTF
                     ISDEVARDLSLQFKRPINPTDQSSYIVDSVAVKNGALHPYFDKAGQKTYERHPLSHFV
                     NLDNLRANNTKGSLPINVIVFDGKSKCDESASKSASVYYSQLMCQPILLLDQALVSDV
                     GDSTEVSVKMFDAYVDTFSATFSVPMEKLKALVATAHSELAKGVALDGVLSTFVSAAR
                     QGVVDTDVDTKDVIECLKLSHHSDLEVTGDSCYNFMLTYNKVENMTPRDLGACIDCNA
                     RHINAQVAKSHNVSLIWNVKDYMSLSEQLRKQIRSAAKKNNIPFRLTCATTRQVVNVI
                     TTKISLKGGKIVSTCFKLMLKATLLCVLAALVCYIVMPVHTLSIHDGYTNEIIGYKAI
                     QDGVTRDIISTDDCFANKHAGFDAWFSQRGGSYKNDKSCPVVAAIITREIGFIVPGLP
                     GTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEYSDFATSACVLAAECTIFKDAMGK
                     PVPYCYDTNLLEGSISYSELRPDTRYVLMDGSIIQFPNTYLEGSVRVVTTFDAEYCRH
                     GTCERSEVGICLSTSGRWVLNNEHYRALSGVFCGVDAMNLIANIFTPLVQPVGALDAS
                     ASVVAGGIIAILVTCAAYYFMKFRRAFGEYNHVVAANALLFLMSFTILCLVPAYSFLP
                     GVYSVFYLYLTFYFTNDVSFLAHLQWFAMFSPIVPFWITAIYVFCISLKHCHWFFNNY
                     LRKRVMFNGVTFSTFEEAALCTFLLNKEMYLKLRSETLLPLTQYNRYLALYNKYKYFS
                     GVLDTTSYREAACCHLAKALNDFSNSGADVLYQPPQTSITSAVLQSGFRKMAFPSGKV
                     EGCMVQVTCGTTTLNGLWLDDTVYCPRHVICTAEDMLNPNYEDLLIRKSNHSFLVQAG
                     NVQLRVIGHSMQNCLLRLKVDTSNPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCA
                     MRPNHTIKGSFLNGSCGSVGFNIDYDCVSFCYMHHMELPTGVHAGTDLEGKFYGPFVD
                     RQTAQAAGTDTTITLNVLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQD
                     HVDILGPLSAQTGIAVLDMCAALKELLQNGMNGRTILGSTILEDEFTPFDVVRQCSGV
                     TFQGKFKKIVKGTHHWMLLTFLTSLLILVQSTQWSLFFFVYENAFLPFTLGIMAIAAC
                     AMLLVKHKHAFLCLFLLPSLATVAYFNMVYMPASWVMRIMTWLELADTSLSGYRLKDC
                     VMYASALVLLILMTARTVYDDAARRVWTLMNVITLVYKVYYGNALDQAISMWALVISA
                     TSNYSGVVTTIMFLARAIVFVCVEYYPLLFITGNTLQCIMLVYCFLGYCCCCYFGLFC
                     LLNRYFRLTLGVYDYLVSTQEFRYMNSQGLLPPKSSIDAFKLNIKLLGIGGKPCIKVA
                     TVQSKMSDVKCTSVVLLSVLQQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMVSL
                     LSVLLSMQGAVDINRLCEEVLDNRATLQAIASEFSSLPSYAAYATAQEAYEQAVANGD
                     SEVVLKKLKKSLNVAKSEFDRDAAMQRKLEKMADQAMTQMYKQARSEDKRAKVTSAMQ
                     TMLFTMLRKLDNDALNNIINNARDGCVPLNIIPLTTAAKLMVVVPDYGTYKNTCDGNT
                     FTYASALWEIQQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANSAAKLQNNELSP
                     VALRQMSCAAGTTQTACTDDNALAYYNNSKGGRFVLALLSDHQDLKWARFPKSDGTGT
                     IYTELEPPCRFVTDTPKGPKMKYLYFIKGLNNLNRGMVLGSLAATVRLQAGNATEVPA
                     NSTVLSFCAFAVDPAKAYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQGS
                     FGGASCCLYCRCHIDHPNPKGFCDLKGKYVQIPTTCANDPVGFTLRNTVCTVCGMWKG
                     YGCSCDQLREPLMQSADASTFFKRVCGVSAARLTPCGTGTSTDVVYRAFDIYNEKVAG
                     FAKFLKTNCCRFQEKDEEGNLLDSYFVVKRHTMSNYQHEETIYNLIKDCPAVAVHDFF
                     KFRVDGDMVPHISRQRLTKYTMADLVYALRHFDEGNCDTLKEILVTYNCCDDDYFNKK
                     DWYDFVENPDILRVYANLGERVRQSLLKTVQFCDAMRDAGIVGVLTLDNQDLNGNWYD
                     FGDFVQVAPGCGVPIVDSYYSLLMPILTLTRALAAESHMDADLAKPLIKWDLLKYDFT
                     EERLCLFDRYFKYWDQTYHPNCINCLDDRCILHCANFNVLFSTVFPPTSFGPLVRKIF
                     VDGVPFVVSTGYHFRELGVVHNQDVNLHSSRLSFKELLVYAADPAMHAASGNLLLDKR
                     TTCFSVAALTNNVAFQTVKPGNFNKDFYDFAVSKGFFKEGSSVELKHFFFAQDGNAAI
                     SDYDYYRYNLPTMCDIRQLLFVVEVVDKYFDCYDGGCINANQVIVNNLDKSAGFPFNK
                     WGKARLYYDSMSYEDQDALFAYTKRNVIPTITQMNLKYAISAKNRARTVAGVSICSTM
                     TNRQFHQKLLKSIAATRGATVVIGTSKFYGGWHNMLKTVYSDVETPHLMGWDYPKCDR
                     AMPNMLRIMASLVLARKHNTCCNLSHRFYRLANECAQVLSEMVMCGGSLYVKPGGTSS
                     GDATTAYANSVFNICQAVTANVNALLSTDGNKIADKYVRNLQHRLYECLYRNRDVDHE
                     FVDEFYAYLRKHFSMMILSDDAVVCYNSNYAAQGLVASIKNFKAVLYYQNNVFMSEAK
                     CWTETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMI
                     ERFVSLAIDAYPLTKHPNQEYADVFHLYLQYIRKLHDELTGHMLDMYSVMLTNDNTSR
                     YWEPEFYEAMYTPHTVLQAVGACVLCNSQTSLRCGACIRRPFLCCKCCYDHVISTSHK
                     LVLSVNPYVCNAPGCDVTDVTQLYLGGMSYYCKSHKPPISFPLCANGQVFGLYKNTCV
                     GSDNVTDFNAIATCDWTNAGDYILANTCTERLKLFAAETLKATEETFKLSYGIATVRE
                     VLSDRELHLSWEVGKPRPPLNRNYVFTGYRVTKNSKVQIGEYTFEKGDYGDAVVYRGT
                     TTYKLNVGDYFVLTSHTVMPLSAPTLVPQEHYVRITGLYPTLNISDEFSSNVANYQKV
                     GMQKYSTLQGPPGTGKSHFAIGLALYYPSARIVYTACSHAAVDALCEKALKYLPIDKC
                     SRIIPARARVECFDKFKVNSTLEQYVFCTVNALPETTADIVVFDEISMATNYDLSVVN
                     ARLRAKHYVYIGDPAQLPAPRTLLTKGTLEPEYFNSVCRLMKTIGPDMFLGTCRRCPA
                     EIVDTVSALVYDNKLKAHKEKSAQCFKMFYKGVITHDVSSAINRPQIGVVREFLTRNP
                     AWRKAVFISPYNSQNAVASKILGLPTQTVDSSQGSEYDYVIFTQTTETAHSCNVNRFN
                     VAITRAKIGILCIMSDRDLYDKLQFTSLEIPRRNVATLQAENVTGLFKDCSKIITGLH
                     PTQAPTHLSVDIKFKTEGLCVDIPGIPKDMTYRRLISMMGFKMNYQVNGYPNMFITRE
                     EAIRHVRAWIGFDVEGCHATRDAVGTNLPLQLGFSTGVNLVAVPTGYVDTENNTEFTR
                     VNAKPPPGDQFKHLIPLMYKGLPWNVVRIKIVQMLSDTLKGLSDRVVFVLWAHGFELT
                     SMKYFVKIGPERTCCLCDKRATCFSTSSDTYACWNHSVGFDYVYNPFMIDVQQWGFTG
                     NLQSNHDQHCQVHGNAHVASCDAIMTRCLAVHECFVKRVDWSVEYPIIGDELRVNSAC
                     RKVQHMVVKSALLADKFPVLHDIGNPKAIKCVPQAEVEWKFYDAQPCSDKAYKIEELF
                     YSYATHHDKFTDGVCLFWNCNVDRYPANAIVCRFDTRVLSNLNLPGCDGGSLYVNKHA
                     FHTPAFDKSAFTNLKQLPFFYYSDSPCESHGKQVVSDIDYVPLKSATCITRCNLGGAV
                     CRHHANEYRQYLDAYNMMISAGFSLWIYKQFDTYNLWNTFTRLQSLENVAYNVVNKGH
                     FDGHAGEAPVSIINNAVYTKVDGIDVEIFESKTTLPVNVAFELWAKRNIKPVPEIKIL
                     NNLGVDIAANTVIWDYKREAPAHVSTIGVCTMTDIAKKPTESACSSLTVLFDGRVEGQ
                     VDLFRNARNGVLITEGSVKGLTPSKGPAQASVNGVTLIGESVKTQFNYFKKVDGIIQQ
                     LPETYFTQSRDLEDFKPRSQMETDFLELAMDEFIQRYKLEGYAFEHIVYGDFSHGQLG
                     GLHLMIGLAKRSQDSPLKLEDFIPMDSTVKNYFITDAQTGSSKCVRSVIDLLLDDFVE
                     IIKSQDLSVISKVVKVTIDYAEISFMLWCKDGHVETFYPKLQASQAWQPGVAMPNLYK
                     MQRMLLEKCDLQNYGENAVIPKGIMMNVAKYTQLCQYLNTLTLAVPYNMRVIHFGAGS
                     DKGVAPGTAVLRQWLPTGTLLVDSDLNDFVSDADSTLIGDCATVHTANKWDLIISDMY
                     DPRTKHVTKENDSKEGFFTYLCGFIKQKLALGGSIAVKITEHSWNADLYKLMGHFSWW
                     TAFVTNVNASSSEAFLIGANYLGKPKEQIDGYTMHANYIFWRNTNPIQLSSYSLFDMS
                     KFPLKLRGTAVMSLKENQINDMIYSLLEKGRLIIRENNRVVVSSDILVNN"
     CDS             20834. .24601
                     /codon_start=1
                     /product="spike"
                     /protein_id="AAR86775.1"
                     /db_xref="GI:40457450"
                     /translation="MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPD
                     EIFRSDTLYLTQDLFLPFYSNVTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRGWV
```

```
                     FGSTMNNKSQSVIIINNSTNVVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCT
                     FEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLKP
                     IFKLPLGINITNFRAILTAFSPAQDIWGTSAAAYFVGYLKPTTFMLKYDENGTITDAV
                     DCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKF
                     PSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGD
                     DVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRP
                     FERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPA
                     TVCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPK
                     TSEILDISPCAFGGVSVITPGTNASSEVAVLYQDVNCTNVSAAIHADQLTPAWRIYST
                     GNNVFQTQAGCLIGAEHVDTSYECDIPIGAGICASYHTVSLLRSTSQKSIVAYTMSLG
                     ADSSIAYSNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGS
                     FCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRS
                     FIEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYT
                     AALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAIS
                     QIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAE
                     VQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYH
                     LMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQ
                     RNFFSPQIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKGELDKYFKNHTSPDV
                     DLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIA
                     GLIAIVMVTILLCCMTSCCSCLKGACSCGSCCKFDEDDSEPVLKGVKLHYT"
     CDS             24610. .25434
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAR86776.1"
                     /db_xref="GI:40457451"
                     /translation="MDLFMRFFTLGSITAQRVKIDNASPASTAHATATIPLQASLPFG
                     WLVIGVAFLAVFQSATKIIALNKRWQLALYKGFQFICNLLLLFVTIYSHLLLVAAGME
                     AQFLYLYALIYFLQCINACRIIMRCWLCWKCKSKNPLLYDANYFVCWHTHNYDYCIPY
                     NSVTDTIVVTEGDGISTPKLKEDYQIGGYSEDRHSGVKDYVVVHGYFTEVYYQLESTQ
                     ITTNTGIENATFFIFNKLVKDPPNVQIHTIDGSSGVANPAMDPIYDEPTTTSVPL"
     CDS             25031. .25495
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAR86777.1"
                     /db_xref="GI:40457452"
                     /translation="MMPTTLFAGTHITMTTVYHITVSQIQLSLLKVTAFQHQNSKKTT
                     KLVVILRIGTQVLKTMSLYMAISPKFTTSLSLHKLLQTLVLKMLHSSSLTSLLKTHRM
                     CKYTQSTALQELLIQQWIQFMMSRRRLLACLCKHKKVSTNLCTHSFRKKQVR"
     CDS             25459. .25689
                     /codon_start=1
                     /product="E"
                     /protein_id="AAR86778.1"

  Query Match             90.9%;  Score 20;  DB 14;  Length 29013;
  Best Local Similarity   75.0%;  Pred. No. 6.5;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        697 GTGAACTCACTCGTGAGCTC 716


RESULT 5
AY394999
LOCUS       AY394999               29350 bp    RNA     linear   VRL 29-JAN-2004
DEFINITION  SARS coronavirus LC2, complete genome.
ACCESSION   AY394999
VERSION     AY394999.1  GI:37624342
KEYWORDS    .
SOURCE      SARS coronavirus LC2
  ORGANISM  SARS coronavirus LC2
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29350)
  AUTHORS   .
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE     From independent foci of epidemic outbreak to large genomic
            alteration in late phase viruses: evolution of the SARS-coronavirus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 29350)
  AUTHORS   .
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE     Direct Submission
  JOURNAL   Submitted (19-SEP-2003) Guangdong, China
FEATURES             Location/Qualifiers
     source          1. .29350
                     /organism="SARS coronavirus LC2"
                     /mol_type="genomic RNA"
                     /isolate="LC2"
                     /db_xref="taxon:249084"
ORIGIN

  Query Match             90.9%;  Score 20;  DB 14;  Length 29350;
  Best Local Similarity   75.0%;  Pred. No. 6.5;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        761 GTGAACTCACTCGTGAGCTC 780


RESULT 6
AY395000
LOCUS       AY395000               29350 bp    RNA     linear   VRL 29-JAN-2004
DEFINITION  SARS coronavirus LC3, complete genome.
ACCESSION   AY395000
VERSION     AY395000.1  GI:37624343
KEYWORDS    .
SOURCE      SARS coronavirus LC3
  ORGANISM  SARS coronavirus LC3
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29350)
  AUTHORS   .
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE     From independent foci of epidemic outbreak to large genomic
            alteration in late phase viruses: evolution of the SARS-coronavirus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 29350)
  AUTHORS   .
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE     Direct Submission
  JOURNAL   Submitted (19-SEP-2003) Guangdong, China
FEATURES             Location/Qualifiers
     source          1. .29350
                     /organism="SARS coronavirus LC3"
                     /mol_type="genomic RNA"
                     /isolate="LC3"
                     /db_xref="taxon:249085"
ORIGIN

  Query Match             90.9%;  Score 20;  DB 14;  Length 29350;
  Best Local Similarity   75.0%;  Pred. No. 6.5;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        761 GTGAACTCACTCGTGAGCTC 780


RESULT 7
AY395001
LOCUS       AY395001               29350 bp    RNA     linear   VRL 29-JAN-2004
DEFINITION  SARS coronavirus LC4, complete genome.
ACCESSION   AY395001
VERSION     AY395001.1  GI:37624344
KEYWORDS    .
SOURCE      SARS coronavirus LC4
  ORGANISM  SARS coronavirus LC4
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29350)
  AUTHORS   .
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE     From independent foci of epidemic outbreak to large genomic
            alteration in late phase viruses: evolution of the SARS-coronavirus
```

```
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 29350)
  AUTHORS   .
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE     Direct Submission
  JOURNAL   Submitted (19-SEP-2003) Guangdong, China
FEATURES             Location/Qualifiers
     source          1. .29350
                     /organism="SARS coronavirus LC4"
                     /mol_type="genomic RNA"
                     /isolate="LC4"
                     /db_xref="taxon:249086"
ORIGIN

  Query Match             90.9%;  Score 20;  DB 14;  Length 29350;
  Best Local Similarity   75.0%;  Pred. No. 6.5;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        761 GTGAACTCACTCGTGAGCTC 780


RESULT 8
AY395002
LOCUS       AY395002               29350 bp    RNA     linear   VRL 29-JAN-2004
DEFINITION  SARS coronavirus LC5, complete genome.
ACCESSION   AY395002
VERSION     AY395002.1  GI:37624345
KEYWORDS    .
SOURCE      SARS coronavirus LC5
  ORGANISM  SARS coronavirus LC5
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29350)
  AUTHORS   .
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE     From independent foci of epidemic outbreak to large genomic
            alteration in late phase viruses: evolution of the SARS-coronavirus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 29350)
  AUTHORS   .
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE     Direct Submission
  JOURNAL   Submitted (19-SEP-2003) Guangdong, China
FEATURES             Location/Qualifiers
     source          1. .29350
                     /organism="SARS coronavirus LC5"
                     /mol_type="genomic RNA"
                     /isolate="LC5"
                     /db_xref="taxon:249087"
ORIGIN

  Query Match             90.9%;  Score 20;  DB 14;  Length 29350;
  Best Local Similarity   75.0%;  Pred. No. 6.5;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        761 GTGAACTCACTCGTGAGCTC 780


RESULT 9
AY394977
LOCUS       AY394977               29433 bp    RNA     linear   VRL 29-JAN-2004
DEFINITION  SARS coronavirus GZ-A, partial genome.
ACCESSION   AY394977
VERSION     AY394977.1  GI:37624320
KEYWORDS    .
SOURCE      SARS coronavirus GZ-A
  ORGANISM  SARS coronavirus GZ-A
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29433)
  AUTHORS   .
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE     From independent foci of epidemic outbreak to large genomic
            alteration in late phase viruses: evolution of the SARS-coronavirus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 29433)
  AUTHORS   .
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE     Direct Submission
  JOURNAL   Submitted (19-SEP-2003) Guangdong, China
FEATURES             Location/Qualifiers
     source          1. .29433
                     /organism="SARS coronavirus GZ-A"
                     /mol_type="genomic RNA"
                     /isolate="GZ-A"
                     /db_xref="taxon:249062"
ORIGIN

  Query Match             90.9%;  Score 20;  DB 14;  Length 29433;
  Best Local Similarity   75.0%;  Pred. No. 6.5;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        725 GTGAACTCACTCGTGAGCTC 744


RESULT 10
AY394985
LOCUS       AY394985               29530 bp    RNA     linear   VRL 29-JAN-2004
DEFINITION  SARS coronavirus HSZ-Bb, complete genome.
ACCESSION   AY394985
VERSION     AY394985.1  GI:37624328
KEYWORDS    .
SOURCE      SARS coronavirus HSZ-Bb
  ORGANISM  SARS coronavirus HSZ-Bb
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29530)
  AUTHORS   .
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE     From independent foci of epidemic outbreak to large genomic
            alteration in late phase viruses: evolution of the SARS-coronavirus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 29530)
  AUTHORS   .
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE     Direct Submission
  JOURNAL   Submitted (19-SEP-2003) Guangdong, China
FEATURES             Location/Qualifiers
     source          1. .29530
                     /organism="SARS coronavirus HSZ-Bb"
                     /mol_type="genomic RNA"
                     /isolate="HSZ-Bb"
                     /db_xref="taxon:249070"
ORIGIN

  Query Match             90.9%;  Score 20;  DB 14;  Length 29530;
  Best Local Similarity   75.0%;  Pred. No. 6.5;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        526 GTGAACTCACTCGTGAGCTC 545


RESULT 11
AY338174
LOCUS       AY338174               29573 bp    RNA     linear   VRL 28-JUL-2003
DEFINITION  SARS coronavirus Taiwan TC1, complete genome.
ACCESSION   AY338174
VERSION     AY338174.1  GI:32493129
```

KEYWORDS .
SOURCE SARS coronavirus Taiwan TC1
ORGANISM SARS coronavirus Taiwan TC1
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales; Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 29573)
AUTHORS Chang,J.-G.C., Lin,T.-H., Chen,C.-M., Lin,C.-S., Chan,W.-L. and Shih,M.-C.
TITLE SARS coronavirus TC1, clinical specimen
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 29573)
AUTHORS Chang,J.-G.C., Lin,T.-H., Chen,C.-M., Lin,C.-S., Chan,W.-L. and Shih,M.-C.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-2003) Department of Molecular Medicine, China Medical University Hospital, 2, Yuh Der Road, Taichung, Taichung 404, Taiwan
REFERENCE 3 (bases 1 to 29573)
AUTHORS Chang,J.-G.C., Lin,T.-H., Chen,C.-M., Lin,C.-S., Chan,W.-L. and Shih,M.-C.
TITLE Direct Submission
JOURNAL Submitted (28-JUL-2003) Department of Molecular Medicine, China Medical University Hospital, 2, Yuh Der Road, Taichung, Taichung 404, Taiwan
REMARK Amino acid sequence updated by submitter
FEATURES Location/Qualifiers
source 1. .29573
/organism="SARS coronavirus Taiwan TC1"
/mol_type="genomic RNA"
/isolate="TC1"
/db_xref="taxon:237639"
/note="clinical specimen"
CDS join(196. .13329,13329. .21416)
/codon_start=1
/product="nonstructural polyprotein pp1ab"
/protein_id="AAQ01594.1"
/db_xref="GI:33285967"
/translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEDYEQN
WNTKHGSGALRELTRELNGGAVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
LDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPFEIKSAKKFDTFKGECPKFVFP
LNSKVKVIQPRVEKKKTEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTC
DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSN
IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
LEILSRERVNINIVGDFHLNEEVAIILASFSASTSAFIDTIKSLDYKSFKTIVESCGN
YKVTKGKPVKGAWNIGQQRSVLTPLCGFPSQAAGVIRSIFARTLDAANHSIPDLQRAA
VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVQQTSQWLSNLLGTTVEKL
RPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKGQIQVASDNIKDCVKCFIDVV
NKALEMCIDQVTIAGAKLRSLNLGEVFIAQSKGLYRQCIRGKEQLQLLMPLKAPKEVT
FLEGDSHDTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTPVCVNGLMLLEIKDKEQ
YCALSPGLLATNNVFRLKGGAPIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNE
KCSVYTVESGTEVTEFACVVAEAVVKTLQPVSDLLTNMGIDLDEWSVATFYLFDDAGE
ENFSSRMYCSFYPPDEEEEDDAECEEEEIDETCEHEYGTEDDYQGLPLEFGASAETVR
VEEEEEEDWLDDTTEQSEIEPEPEPTPEEPVNQFTGYLKLTDNVAIKCVDIVKEAQSA
NPMVIVNAANIHLKHGGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL
AKKCLHVVGPNLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQT
VRTQVYIAVNDKALYEQVVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPVDV
KPKIKACIDEVTTTLEETKFLTNKLLLFADINGKLYHDSQNMLRGEDMSFLEKDAPYM
VGDVITSGDITCVVIPSKKAGGTTEMLSRALKKVPVDEYITTYPGQGCAGYTLEEAKT
ALKKCKSAFYVLPSEAPNAKEEILGTVSWNLREMLAHAEETRKLMPICMDVRAIMATI
QRKYKGIKIQEGIVDYGVRFFFYTSKEPVASIITKLNSLNEPLVTMPIGYVTHGFNLE
EAARCMRSLKAPAVVSVSSPDAVTTYNGYLTSSSKTSEEHFVETVSLAGSYRDWSYSG
QRTELGVEFLKRGDKIVYHTLESPVEFHLDGEVLSLDKLKSLLSLREVKTIKVFTTVD
NTNLHTQLVDMSMTYGQQFGPTYLDGADVTKIKPHVNHEGKTFFVLPSDDTLRSEAFE
YYHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVLLALQQLEVKFN
APALQEAYYRARAGDAANFCALILAYSNKTVGELGDVRETMTHLLQHANLESAKRVLN
VVCKHCGQKTTTLTGVEAVMYMGTLSYDNLKTGVSIPCVCGRDATQYLVQQESSFVMM
SAPPAEYKLQQGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTD
VFYKETSYTTTIKPVSYKLDGVTYTEIEPKLDGYYKKDNAYYTEQPIDLVPTQPLPNA
SFDNFKLTCSNTKFADDLNQMTGFTKPASRELSVTFFPDLNGDVVAIDYRHYSASFKK
GAKLLHKPIVWHINQATTKTTFKPNTWCLRCLWSTKPVDTSNSFEVLAVEDTQGMDNL
ACESQQPTSEEVVENPTIQKEVIECDVKTTEVVGNVILKPSDEGVKVTQELGHEDLMA
AYVENTSITIKKPNELSLALGLKTIATHGIAAINSVPWSKILAYVKPFLGQAAITTSN
CAKRLAQRVFNNYMPYVFTLLFQLCTFTKSTNSRIRASLPTTIAKNSVKSVAKLCLDA
GINYVKSPKFSKLFTIAMWLLLLSICLGSLICVTAAFGVLLSNFGAPSYCNGVRELYL
NSSNVTTMDFCEGSFPCSICLSGLDSLDSYPALETIQVTISSYKLDLTILGLAAEWVL
AYMLFTKFFYLLGLSAIMQVFFGYFASHFISNSWLMWFIISIVQMAPVSAMVRMYIFF
ASFYYIWKSYVHIMDGCTSSTCMMCYKRNRATRVECTTIVNGMKRSFYVYANGGRGFC
KTHNWNCLNCDTFCTGSTFISDEVARDLSLQFKRPINPTDQSSYIVDSVAVKNGALHL
YFDKAGQKTYERHPLSHFVNLDNLRANNTKGSLPINVIVFDGKSKCDESASKSASVYY
SQLMCQPILLLDQALVSDVGDSTEVSVKMFDAYVDTFSATFSVPMEKLKALVATAHSE
LAKGVALDGVLSTFVSAARQGVVDTDVDTKDVIECLKLSHHSDLEVTGDSCNNFMLTY
NKVENMTPRDLGACIDCNARHINAQVAKSHNVSLIWNVKDYMSLSEQLRKQIRSAAKK
NNIPFRLTCATTRQVVNVITTKISLKGGKIVSTCFKLMLKATLLCVLAALVCYIVMPV
HTLSIHDGYTNEIIGYKAIQDGVTRDIISTDDCFANKHAGFDAWFSQRGGSYKNDKSC
PVVAAIITREIGFIVPGLPGTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEYSDFA
TSACVLAAECTIFKDAMGKPVPYCYDTNLLEGSISYSELRPDTRYVLMDGSIIQFPNT
YLEGSVRVVTTFDAEYCRHGTCERSEVGICLSTSGRWVLNNEHYRALSGVFCGVDAMN
LIANIFTPLVQPVGALDVSASVVAGGIIAILVTCAAYYFMKFRRVFGEYNHVVAANAL
LFLMSFTILCLVPAYSFLPGVYSVFYLYLTFYFTNDVSFLAHLQWFAMFSPIVPFWIT
AIYVFCISLKHCHWFFNNYLRKRVMFNGVTFSTFEEAALCTFLLNKEMYLKLRSETLL
PLTQYNRYLALYNKYKYFSGALDTTSYREAACCHLAKALNDFSNSGADVLYQPPQTSI
TSAVLQSGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDTVYCPRHVICTAEDMLNP
NYEDLLIRKSNHSFLVQAGNVQLRVIGHSMQNCLLRLKVDTSNPKTPKYKFVRIQPGQ
TFSVLACYNGSPSGVYQCAMRPNHTIKGSFLNGSCGSVGFNIDYDCVSFCYMHHMELP
TGVHAGTDLEGKFYGPFVDRQTAQAAGTDTTITLNVLAWLYAAVINGDRWFLNRFTTT
LNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCAALKELLQNGMNGRTILGS
TILEDEFTPFDVVRQCSGVTFQGKFKKIVKGTHHWMLLTFLTSLLILVQSTQWSLFFF
VYENAFLPFTLGIMAIAACAMLLVKHKHAFLCLFLLPSLATVAYFNMVYMPASWVMRI
MTWLELADTSLSGYRLKDCVMYASALVLLILMTARTVYDDAARRVWTLMNVITLVYKV
YYGNALDQAISMWALVISVTSNYSGVVTTIMFLARAIVFVCVEYYPLLFITGNTLQCI
MLVYCFLGYCCCCYFGLFCLLNRYFRLTLGVYDYLVSTQEFRYMNSQGLLPPKSSIDA
FKLNIKLLGIGGKPCIKVATVQSKMSDVKCTSVVLLSVLQQLRVESSSKLWAQCVQLH
NDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINRLCEEMLDNRATLQAIASEFSSLPS
YAAYATAQEAYEQAVANGDSEVVLKKLKKSLNVAKSEFDRDAAMQRKLEKMADQAMTQ
MYKQARSEDKRAKVTSAMQTMLFTMLRKLDNDALNNIINNARDGCVPLNIIPLTTAAK
LMVVVPDYGTYKNTCDGNTFTYASALWEIQQVVDADSKIVQLSEINMDNSPNLAWPLI
VTALRANSAVKLQNNELSPVALRQMSCAAGTTQTACTDDNALAYYNNSKGGRFVLALL
SDHQDLKWARFPKSDGTGTIYTELEPPCRFVTDTPKGPKVKYLYFIKGLNNLNRGMVL
GSLAATVRLQAGNATEVPANSTVLSFCAFAVDPAKAYKDYLASGGQPITNCVKMLCTH
TGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHPNPKGFCDLKGKYVQIPTTCAND
PVGFTLRNTVCTVCGMWKGYGCSCDQLREPLMQSADASTFLNRVCGVSAARLTPCGTG
TSTDVVYRAFDIYNEKVAGFAKFLKTNCCRFQEKDEEGNLLDSYFVVKRHTMSNYQHE
ETIYNLVKDCPAVAVHDFFKFRVDGDMVPHISRQRLTKYTMADLVYALRHFDEGNCDT
LKEILVTYNCCDDDYFNKKDWYDFVENPDILRVYANLGERVRQSLLKTVQFCDAMRDA
GIVGVLTLDNQDLNGNWYDFGDFVQVAPGCGVPIVDSYYSLLMPILTLTRALAAESHM
DADLAKPLIKWDLLKYDFTEERLCLFDRYFKYWDQTYHPNCINCLDDRCILHCANFNV
LFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVNLHSSRLSFKELLV
YAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFYDFAVSKGFFKE
GSSVELKHFFFAQDGNAAISDYDYYRYNLPTMCDIRQLLFVVEVVDKYFDCYDGGCIN
ANQVIVNNLDKSAGFPFNKWGKARLYYDSMSYEDQDALFAYTKRNVIPTITQMNLKYA
ISAKNRARTVAGVSICSTMTNRQFHQKLLKSIAATRGATVVIGTSKFYGGWHNMLKTV
YSDVETPHLMGWDYPKCDRAMPNMLRIMASLVLARKHNTCCNLSHRFYRLANECAQVL
SEMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQAVTANVNALLSTDGNKIADKYVR
NLQHRLYECLYRNRDVDHEFVDEFYAYLRKHFSMMILSDDAVVCYNSNYAAQGLVASI
KNFKAVLYYQNNVFMSEAKCWTETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRI
LGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADVFHLYLQYIRKLHDEL
TGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTVLQAVGACVLCNSQTSLRCGACIR
RPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCDVTDVTQLYLGGMSYYCKSHKPPI
SFPLCANGQVFGLYKNTCVGSDNVTDFNAIATCDWTNAGDYILANTCTERLKLFAAET
LKATEETFKLSYGIATVREVLSDRELHLSWEVGKPRPPLNRNYVFTGYRVTKNSKVQI
GEYTFEKGDYGDAVVYRGTTTYKLNVGDYFVLTSHTVMPLSAPTLVPQEHYVRITGLY
PTLNISDEFSSNVANYQKVGMQKYSTLQGPPGTGKSHFAIGLALYYPSARIVYTACSH
AAVDALCEKALKYLPIDKCSRIIPARARVECFDKFKVNSTLEQYVFCTVNALPETTAD
IVVFDEISMATNYDLSVVNARLRAKHYVYIGDPAQLPAPRTLLTKGTLEPEYFNSVCR
LMKTIGPDMFLGTCRRCPAEIVDTVSALVYDNKLKAHKDKSAQCFKMFYKGVITHDVS
SAINRPQIGVVREFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQTVDSSQGSEYDY
VIFTQTTETAHSCNVNRFNVAITRAKIGILCIMSDRDLYDKLQFTSLEIPRRNVATLQ
AENVTGLFKDCSKIITGLHPTQAPTHLSVDIKFKTEGLCVDIPGIPKDMTYRRLISMM
GFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEGCHATRDAVGTNLPLQLGFSTGVN
LVAVPTGYVDTENNTEFTRVNAKPPPGDQFKHLIPLMYKGLPWNVVRIKIVQMLSDTL
KGLSDRVVFVLWAHGFELTSMKYFVKIGPERTCCLCDKRATCFSTSSDTYACWNHSVG
FDYVYNPFMIDVQQWGFTGNLQSNHDQHCQVHGNAHVASCDAIMTRCLAVHECFVKRV
DWSVEYPIIGDELRVNSACRKVQHMVVKSALLADKFPVLHDIGNPKAIKCVPQAEVEW
KFYDAQPCSDKAYKIEELFYSYATHHDKFTDGVCLFWNCNVDRYPANAIVCRFDTRVL

SNLNLPGCDGGSLYVNKHAFHTPAFDKSAFTNLKQLPFFYYSDSPCESHGKQVVSDID
YVPLKSATCITRCNLGGAVCRHHANEYRQYLDAYNMMISAGFSLWIYKQFDTYNLWNT
FTRLQSLENVAYNVVNKGHFDGHAGEAPVSIINNAVYTKVDGIDVEIFENKTTLPVNV
AFELWAKRNIKPVPEIKILNNLGVDIAANTVIWDYKREAPAHVSTIGVCTMTDIAKKP
TESACSSLTVLFDGRVEGQVDLFRNARNGVLITEGSVKGLTPSKGPAQASVNGVTLIG
ESVKTQFNYFKKVDGIIQQLPETYFTQSRDLEDFKPRSQMETDFLELAMDEFIQRYKL
EGYAFEHIVYGDFSHGQLGGLHLMIGLAKRSQDSPLKLEDFIPMDSTVKNYFITDAQT
GSSKCVCSVIDLLLDDFVEIIKSQDLSVISKVVKVTIDYAEISFMLWCKDGHVETFYP
KLQASQAWQPGVAMPNLYKMQRMLLEKCDLQNYGENAVIPKGIMMNVAKYTQLCQYLN
TLTLAVPYNMRVIHFGAGSDKGVAPGTAVLRQWLPTGTLLVDSDLNDFVSDADSTLIG
DCATVHTANKWDLIISDMYDPRTKHVTKENDSKEGFFTYLCGFIKQKLALGGSIAVKI
TEHSWNADLYKLMGHFSWWTAFVTNVNASSSEAFLIGANYLGKPKEQIDGYTMHANYI
FWRNTNPIQLSSYSLFDMSKFPLKLRGTAVMSLKENQINDMIYSLLEKGRLIIRENNR
VVVSSDILVNN"
CDS 196. .13344
/codon_start=1
/product="nonstructural polyprotein pp1a"
/protein_id="AAQ01595.1"
/db_xref="GI:33285968"
/translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEDYEQN
WNTKHGSGALRELTRELNGGAVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
LDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPFEIKSAKKFDTFKGECPKFVFP
LNSKVKVIQPRVEKKKTEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTC
DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSN
IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
LEILSRERVNINIVGDFHLNEEVAIILASFSASTSAFIDTIKSLDYKSFKTIVESCGN
YKVTKGKPVKGAWNIGQQRSVLTPLCGFPSQAAGVIRSIFARTLDAANHSIPDLQRAA
VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVQQTSQWLSNLLGTTVEKL
RPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKGQIQVASDNIKDCVKCFIDVV
NKALEMCIDQVTIAGAKLRSLNLGEVFIAQSKGLYRQCIRGKEQLQLLMPLKAPKEVT
FLEGDSHDTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTPVCVNGLMLLEIKDKEQ
YCALSPGLLATNNVFRLKGGAPIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNE
KCSVYTVESGTEVTEFACVVAEAVVKTLQPVSDLLTNMGIDLDEWSVATFYLFDDAGE
ENFSSRMYCSFYPPDEEEEDDAECEEEEIDETCEHEYGTEDDYQGLPLEFGASAETVR
VEEEEEEDWLDDTTEQSEIEPEPEPTPEEPVNQFTGYLKLTDNVAIKCVDIVKEAQSA
NPMVIVNAANIHLKHGGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL
AKKCLHVVGPNLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQT
VRTQVYIAVNDKALYEQVVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPVDV
KPKIKACIDEVTTTLEETKFLTNKLLLFADINGKLYHDSQNMLRGEDMSFLEKDAPYM
VGDVITSGDITCVVIPSKKAGGTTEMLSRALKKVPVDEYITTYPGQGCAGYTLEEAKT
ALKKCKSAFYVLPSEAPNAKEEILGTVSWNLREMLAHAEETRKLMPICMDVRAIMATI
QRKYKGIKIQEGIVDYGVRFFFYTSKEPVASIITKLNSLNEPLVTMPIGYVTHGFNLE
EAARCMRSLKAPAVVSVSSPDAVTTYNGYLTSSSKTSEEHFVETVSLAGSYRDWSYSG
QRTELGVEFLKRGDKIVYHTLESPVEFHLDGEVLSLDKLKSLLSLREVKTIKVFTTVD
NTNLHTQLVDMSMTYGQQFGPTYLDGADVTKIKPHVNHEGKTFFVLPSDDTLRSEAFE
YYHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVLLALQQLEVKFN
APALQEAYYRARAGDAANFCALILAYSNKTVGELGDVRETMTHLLQHANLESAKRVLN
VVCKHCGQKTTTLTGVEAVMYMGTLSYDNLKTGVSIPCVCGRDATQYLVQQESSFVMM

Query Match 90.9%; Score 20; DB 14; Length 29573;
Best Local Similarity 75.0%; Pred. No. 6.5;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCGUGAGCUC 20
|:||||:|||:||:||||:|
Db 707 GTGAACTCACTCGTGAGCTC 726

RESULT 12
AY338175
LOCUS AY338175 29573 bp RNA linear VRL 28-JUL-2003
DEFINITION SARS coronavirus Taiwan TC2, complete genome.
ACCESSION AY338175
VERSION AY338175.1 GI:32493130
KEYWORDS .
SOURCE SARS coronavirus Taiwan TC2
ORGANISM SARS coronavirus Taiwan TC2
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 29573)
AUTHORS Chang,J., Lin,T., Chen,C., Lin,T., Chan,W. and Shih,M.
TITLE SARS coronavirus TC2, clinical specimen
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 29573)
AUTHORS Chang,J., Lin,T., Chen,C., Lin,T., Chan,W. and Shih,M.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2003) Department of Molecular Medicine, China Medical University Hospital, 2, Yuh Der Road, Taichung, Taichung 404, Taiwan
REFERENCE 3 (bases 1 to 29573)
AUTHORS Chang,J., Lin,T., Chen,C., Lin,T., Chan,W. and Shih,M.
TITLE Direct Submission
JOURNAL Submitted (28-JUL-2003) Department of Molecular Medicine, China Medical University Hospital, 2, Yuh Der Road, Taichung, Taichung 404, Taiwan
REMARK Amino acid sequence updated by submitter
FEATURES Location/Qualifiers
source 1. .29573
/organism="SARS coronavirus Taiwan TC2"
/mol_type="genomic RNA"
/isolate="TC2"
/db_xref="taxon:237640"
/note="clinical specimen"
CDS join(196. .13329,13329. .21416)
/codon_start=1
/product="nonstructural polyprotein pp1ab"
/protein_id="AAQ01606.1"
/db_xref="GI:33285979"
/translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEDYEQN
WNTKHGSGALRELTRELNGGAVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
LDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPFEIKSAKKFDTFKGECPKFVFP
LNSKVKVIQPRVEKKKTEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTC
DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSN
IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
LEILSRERVNINIVGDFHLNEEVAIILASFSASTSAFIDTIKSLDYKSFKTIVESCGN
YKVTKGKPVKGAWNIGQQRSVLTPLCGFPSQAAGVIRSIFARTLDAANHSIPDLQRAA
VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVQQTSQWLSNLLGTTVEKL
RPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKGQIQVASDNIKDCVKCFIDVV
NKALEMCIDQVTIAGAKLRSLNLGEVFIAQSKGLYRQCIRGKEQLQLLMPLKAPKEVT
FLEGDSHDTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTPVCVNGLMLLEIKDKEQ
YCALSPGLLATNNVFRLKGGAPIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNE
KCSVYTVESGTEVTEFACVVAEAVVKTLQPVSDLLTNMGIDLDEWSVATFYLFDDAGE
ENFSSRMYCSFYPPDEEEEDDAECEEEEIDETCEHEYGTEDDYQGLPLEFGASAETVR
VEEEEEEDWLDDTTEQSEIEPEPEPTPEEPVNQFTGYLKLTDNVAIKCVDIVKEAQSA
NPMVIVNAANIHLKHGGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL
AKKCLHVVGPNLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQT
VRTQVYIAVNDKALYEQVVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPVDV
KPKIKACIDEVTTTLEETKFLTNKLLLFADINGKLYHDSQNMLRGEDMSFLEKDAPYM
VGDVITSGDITCVVIPSKKAGGTTEMLSRALKKVPVDEYITTYPGQGCAGYTLEEAKT
ALKKCKSAFYVLPSEAPNAKEEILGTVSWNLREMLAHAEETRKLMPICMDVRAIMATI
QRKYKGIKIQEGIVDYGVRFFFYTSKEPVASIITKLNSLNEPLVTMPIGYVTHGFNLE
EAARCMRSLKAPAVVSVSSPDAVTTYNGYLTSSSKTSEEHFVETVSLAGSYRDWSYSG
QRTELGVEFLKRGDKIVYHTLESPVEFHLDGEVLSLDKLKSLLSLREVKTIKVFTTVD
NTNLHTQLVDMSMTYGQQFGPTYLDGADVTKIKPHVNHEGKTFFVLPSDDTLRSEAFE
YYHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVLLALQQLEVKFN
APALQEAYYRARAGDAANFCALILAYSNKTVGELGDVRETMTHLLQHANLESAKRVLN
VVCKHCGQKTTTLTGVEAVMYMGTLSYDNLKTGVSIPCVCGRDATQYLVQQESSFVMM
SAPPAEYKLQQGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTD
VFYKETSYTTTIKPVSYKLDGVTYTEIEPKLDGYYKKDNAYYTEQPIDLVPTQPLPNA
SFDNFKLTCSNTKFADDLNQMTGFTKPASRELSVTFFPDLNGDVVAIDYRHYSASFKK
GAKLLHKPIVWHINQATTKTTFKPNTWCLRCLWSTKPVDTSNSFEVLAVEDTQGMDNL
ACESQQPTSEEVVENPTIQKEVIECDVKTTEVVGNVILKPSDEGVKVTQELGHEDLMA
AYVENTSITIKKPNELSLALGLKTIATHGIAAINSVPWSKILAYVKPFLGQAAITTSN
CAKRLAQRVFNNYMPYVFTLLFQLCTFTKSTNSRIRASLPTTIAKNSVKSVAKLCLDA
GINYVKSPKFSKLFTIAMWLLLLSICLGSLICVTAAFGVLLSNFGAPSYCNGVRELYL
NSSNVTTMDFCEGSFPCSICLSGLDSLDSYPALETIQVTISSYKLDLTILGLAAEWVL
AYMLFTKFFYLLGLSAIMQVFFGYFASHFISNSWLMWFIISIVQMAPVSAMVRMYIFF
ASFYYIWKSYVHIMDGCTSSTCMMCYKRNRATRVECTTIVNGMKRSFYVYANGGRGFC
KTHNWNCLNCDTFCTGSTFISDEVARDLSLQFKRPINPTDQSSYIVDSVAVKNGALHL
YFDKAGQKTYERHPLSHFVNLDNLRANNTKGSLPINVIVFDGKSKCDESASKSASVYY
SQLMCQPILLLDQVLVSDVGDSTEVSVKMFDAYVDTFSATFSVPMEKLKALVATAHSE
LAKGVALDGVLSTFVSAARQGVVDTDVDTKDVIECLKLSHHSDLEVTGDSCNNFMLTY

```
NKVENMTPRDLGACIDCNARHINAQVAKSHNVSLIWNVKDYMSLSEQLRKQIRSAAKK
NNIPFRLTCATTRQVVNVITTKISLKGGKIVSTCFKLMLKATLLCVLAALVCYIVMPV
HTLSIHDGYTNEIIGYKAIQDGVTRDIISTDDCFANKHAGFDAWFSQRGGSYKNDKSC
PVVAAIITREIGFIVPGLPGTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEYSDFA
TSACVLAAECTIFKDAMGKPVPYCYDTNLLEGSISYSELRPDTRYVLMDGSIIQFPNT
YLEGSVRVVTTFDAEYCRHGTCERSEVGICLSTSGRWVLNNEHYRALSGVFCGVDAMN
LIANIFTPLVQPVGALDVSASVVAGGIIAILVTCAAYYFMKFRRVFGEYNHVVAANAL
LFLMSFTILCLVPAYSFLPGVYSVFYLYLTFYFTNDVSFLAHLQWFAMFSPIVPFWIT
AIYVFCISLKHCHWFFNNYLRKRVMFNGVTFSTFEEAALCTFLLNKEMYLKLRSETLL
PLTQYNRYLALYNKYKYFSGALDTTSYREAACCHLAKALNDFSNSGADVLYQPPQTSI
TSAVLQSGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDTVYCPRHVICTAEDMLNP
NYEDLLIRKSNHSFLVQAGNVQLRVIGHSMQNCLLRLKVDTSNPKTPKYKFVRIQPGQ
TFSVLACYNGSPSGVYQCAMRPNHTIKGSFLNGSCGSVGFNIDYDCVSFCYMHHMELP
TGVHAGTDLEGKFYGPFVDRQTAQAAGTDTTITLNVLAWLYAAVINGDRWFLNRFTTT
LNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCAALKELLQNGMNGRTILGS
TILEDEFTPFDVVRQCSGVTFQGKFKKIVKGTHHWMLLTFLTSLLILVQSTQWSLFFF
VYENAFLPFTLGIMAIAACAMLLVKHKHAFLCLFLLPSLATVAYFNMVYMPASWVMRI
MTWLELADTSLSGYRLKDCVMYASALVLLILMTARTVYDDAARRVWTLMNVITLVYKV
YYGNALDQAISMWALVISVTSNYSGVVTTIMFLARAIVFVCVEYYPLLFITGNTLQCI
MLVYCFLGYCCCCYFGLFCLLNRYFRLTLGVYDYLVSTQEFRYMNSQGLLPPKSSIDA
FKLNIKLLGIGGKPCIKVATVQSKMSDVKCTSVVLLSVLQQLRVESSSKLWAQCVQLH
NDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINRLCEEMLDNRATLQAIASEFSSLPS
YAAYATAQEAYEQAVANGDSEVVLKKLKKSLNVAKSEFDRDAAMQRKLEKMADQAMTQ
MYKQARSEDKRAKVTSAMQTMLFTMLRKLDNDALNNIINNARDGCVPLNIIPLTTAAK
LMVVVPDYGTYKNTCDGNTFTYASALWEIQQVVDADSKIVQLSEINMDNSPNLAWPLI
VTALRANSAVKLQNNELSPVALRQMSCAAGTTQTACTDDNALAYYNNSKGGRFVLALL
SDHQDLKWARFPKSDGTGTIYTELEPPCRFVTDTPKGPKVKYLYFIKGLNNLNRGMVL
GSLAATVRLQAGNATEVPANSTVLSFCAFAVDPAKAYKDYLASGGQPITNCVKMLCTH
TGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHPNPKGFCDLKGKYVQIPTTCAND
PVGFTLRNTVCTVCGMWKGYGCSCDQLREPLMQSADASTFLNRVCGVSAARLTPCGTG
TSTDVVYRAFDIYNEKVAGFAKFLKTNCCRFQEKDEEGNLLDSYFVVKRHTMSNYQHE
ETIYNLVKDCPAVAVHDFFKFRVDGDMVPHISRQRLTKYTMADLVYALRHFDEGNCDT
LKEILVTYNCCDDDYFNKKDWYDFVENPDILRVYANLGERVRQSLLKTVQFCDAMRDA
GIVGVLTLDNQDLNGNWYDFGDFVQVAPGCGVPIVDSYYSLLMPILTLTRALAAESHM
DADLAKPLIKWDLLKYDFTEERLCLFDRYFKYWDQTYHPNCINCLDDRCILHCANFNV
LFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVNLHSSRLSFKELLV
YAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFYDFAVSKGFFKE
GSSVELKHFFFAQDGNAAISDYDYYRYNLPTMCDIRQLLFVVEVVDKYFDCYDGGCIN
ANQVIVNNLDKSAGFPFNKWGKARLYYDSMSYEDQDALFAYTKRNVIPTITQMNLKYA
ISAKNRARTVAGVSICSTMTNRQFHQKLLKSIAATRGATVVIGTSKFYGGWHNMLKTV
YSDVETPHLMGWDYPKCDRAMPNMLRIMASLVLARKHNTCCNLSHRFYRLANECAQVL
SEMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQAVTANVNALLSTDGNKIADKYVR
NLQHRLYECLYRNRDVDHEFVDEFYAYLRKHFSMMILSDDAVVCYNSNYAAQGLVASI
KNFKAVLYYQNNVFMSEAKCWTETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRI
LGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADVFHLYLQYIRKLHDEL
TGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTVLQAVGACVLCNSQTSLRCGACIR
RPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCDVTDVTQLYLGGMSYYCKSHKPPI
SFPLCANGQVFGLYKNTCVGSDNVTDFNAIATCDWTNAGDYILANTCTERLKLFAAET
LKATEETFKLSYGIATVREVLSDRELHLSWEVGKPRPPLNRNYVFTGYRVTKNSKVQI
GEYTFEKGDYGDAVVYRGTTTYKLNVGDYFVLTSHTVMPLSAPTLVPQEHYVRITGLY
PTLNISDEFSSNVANYQKVGMQKYSTLQGPPGTGKSHFAIGLALYYPSARIVYTACSH
AAVDALCEKALKYLPIDKCSRIIPARARVECFDKFKVNSTLEQYVFCTVNALPETTAD
IVVFDEISMATNYDLSVVNARLRAKHYVYIGDPAQLPAPRTLLTKGTLEPEYFNSVCR
LMKTIGPDMFLGTCRRCPAEIVDTVSALVYDNKLKAHKDKSAQCFKMFYKGVITHDVS
SAINRPQIGVVREFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQTVDSSQGSEYDY
VIFTQTTETAHSCNVNRFNVAITRAKIGILCIMSDRDLYDKLQFTSLEIPRRNVATLQ
AENVTGLFKDCSKIITGLHPTQAPTHLSVDIKFKTEGLCVDIPGIPKDMTYRRLISMM
GFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEGCHATRDAVGTNLPLQLGFSTGVN
LVAVPTGYVDTENNTEFTRVNAKPPPGDQFKHLIPLMYKGLPWNVVRIKIVQMLSDTL
KGLSDRVVFVLWAHGFELTSMKYFVKIGPERTCCLCDKRATCFSTSSDTYACWNHSVG
FDYVYNPFMIDVQQWGFTGNLQSNHDQHCQVHGNAHVASCDAIMTRCLAVHECFVKRV
DWSVEYPIIGDELRVNSACRKVQHMVVKSALLADKFPVLHDIGNPKAIKCVPQAEVEW
KFYDAQPCSDKAYKIEELFYSYATHHDKFTDGVCLFWNCNVDRYPANAIVCRFDTRVL
SNLNLPGCDGGSLYVNKHAFHTPAFDKSAFTNLKQLPFFYYSDSPCESHGKQVVSDID
YVPLKSATCITRCNLGGAVCRHHANEYRQYLDAYNMMISAGFSLWIYKQFDTYNLWNT
FTRLQSLENVAYNVVNKGHFDGHAGEAPVSIINNAVYTKVDGIDVEIFENKTTLPVNV
AFELWAKRNIKPVPEIKILNNLGVDIAANTVIWDYKREAPAHVSTIGVCTMTDIAKKP
TESACSSLTVLFDGRVEGQVDLFRNARNGVLITEGSVKGLTPSKGPAQASVNGVTLIG
ESVKTQFNYFKKVDGIIQQLPETYFTQSRDLEDFKPRSQMETDFLELAMDEFIQRYKL
EGYAFEHIVYGDFSHGQLGGLHLMIGLAKRSQDSPLKLEDFIPMDSTVKNYFITDAQT
GSSKCVCSVIDLLLDDFVEIIKSQDLSVISKVVKVTIDYAEISFMLWCKDGHVETFYP
KLQASQAWQPGVAMPNLYKMQRMLLEKCDLQNYGENAVIPKGIMMNVAKYTQLCQYLN
TLTLAVPYNMRVIHFGAGSDKGVAPGTAVLRQWLPTGTLLVDSDLNDFVSDADSTLIG
```

```
DCATVHTANKWDLIISDMYDPRTKHVTKENDSKEGFFTYLCGFIKQKLALGGSIAVKI
TEHSWNADLYKLMGHFSWWTAFVTNVNASSSEAFLIGANYLGKPKEQIDGYTMHANYI
FWRNTNPIQLSSYSLFDMSKFPLKLRGTAVMSLKENQINDMIYSLLEKGRLIIRENNR
VVVSSDILVNN"
     CDS             196. .13344
                     /codon_start=1
                     /product="nonstructural polyprotein ppla"
                     /protein_id="AAQ01607.1"
                     /db_xref="GI:33285980"
                     /translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
                     HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
                     TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEDYEQN
                     WNTKHGSGALRELTRELNGGAVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
                     LDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPFEIKSAKKFDTFKGECPKFVFP
                     LNSKVKVIQPRVEKKKTEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTC
                     DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSN
                     IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
                     LEILSRERVNINIVGDFHLNEEVAIILASFSASTSAFIDTIKSLDYKSFKTIVESCGN
                     YKVTKGKPVKGAWNIGQQRSVLTPLCGFPSQAAGVIRSIFARTLDAANHSIPDLQRAA
                     VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVQQTSQWLSNLLGTTVEKL
                     RPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKGQIQVASDNIKDCVKCFIDVV
                     NKALEMCIDQVTIAGAKLRSLNLGEVFIAQSKGLYRQCIRGKEQLQLLMPLKAPKEVT
                     FLEGDSHDTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTPVCVNGLMLLEIKDKEQ
                     YCALSPGLLATNNVFRLKGGAPIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNE
                     KCSVYTVESGTEVTEFACVVAEAVVKTLQPVSDLLTNMGIDLDEWSVATFYLFDDAGE
                     ENFSSRMYCSFYPPDEEEEDDAECEEEEIDETCEHEYGTEDDYQGLPLEFGASAETVR
                     VEEEEEEDWLDDTTEQSEIEPEPEPTPEEPVNQFTGYLKLTDNVAIKCVDIVKEAQSA
                     NPMVIVNAANIHLKHGGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL
                     AKKCLHVVGPNLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQT
                     VRTQVYIAVNDKALYEQVVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPVDV
                     KPKIKACIDEVTTTLEETKFLTNKLLLFADINGKLYHDSQNMLRGEDMSFLEKDAPYM
                     VGDVITSGDITCVVIPSKKAGGTTEMLSRALKKVPVDEYITTYPGQGCAGYTLEEAKT
                     ALKKCKSAFYVLPSEAPNAKEEILGTVSWNLREMLAHAEETRKLMPICMDVRAIMATI
                     QRKYKGIKIQEGIVDYGVRFFFYTSKEPVASIITKLNSLNEPLVTMPIGYVTHGFNLE
                     EAARCMRSLKAPAVVSVSSPDAVTTYNGYLTSSSKTSEEHFVETVSLAGSYRDWSYSG
                     QRTELGVEFLKRGDKIVYHTLESPVEFHLDGEVLSLDKLKSLLSLREVKTIKVFTTVD
                     NTNLHTQLVDMSMTYGQQFGPTYLDGADVTKIKPHVNHEGKTFFVLPSDDTLRSEAFE
                     YYHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVLLALQQLEVKFN
                     APALQEAYYRARAGDAANFCALILAYSNKTVGELGDVRETMTHLLQHANLESAKRVLN
                     VVCKHCGQKTTTLTGVEAVMYMGTLSYDNLKTGVSIPCVCGRDATQYLVQQESSFVMM
                     SAPPAEYKLQQGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTD
                     VFYKETSYTTTIKPVSYKLDGVTYTEIEPKLDGYYKKDNAYYTEQPIDLVPTQPLPNA
                     SFDNFKLTCSNTKFADDLNQMTGFTKPASRELSVTFFPDLNGDVVAIDYRHYSASFKK

  Query Match             90.9%;  Score 20;  DB 14;  Length 29573;
  Best Local Similarity   75.0%;  Pred. No. 6.5;
  Matches    15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        707 GTGAACTCACTCGTGAGCTC 726


RESULT 13
AY348314
LOCUS       AY348314               29573 bp    RNA     linear   VRL 29-JUL-2003
DEFINITION  SARS coronavirus Taiwan TC3, complete genome.
ACCESSION   AY348314
VERSION     AY348314.1  GI:33188324
KEYWORDS    .
SOURCE      SARS coronavirus Taiwan TC3
  ORGANISM  SARS coronavirus Taiwan TC3
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29573)
  AUTHORS   Chang,J.-G., Lin,T.-H., Chen,C.-M., Lin,T.-S., Chan,W.-L. and
            Shih,M.-C.
  TITLE     SARS coronavirus TC3 clinical specimen
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 29573)
  AUTHORS   Chang,J.-G., Lin,T.-H., Chen,C.-M., Lin,T.-S., Chan,W.-L. and
            Shih,M.-C.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (23-JUL-2003) Department of Molecular Medicine, China
            Medical University Hospital, 2, Yuh Der Road, Taichung, Taichung
            404, Taiwan
FEATURES             Location/Qualifiers
     source          1. .29573
                     /organism="SARS coronavirus Taiwan TC3"
                     /mol_type="genomic RNA"
                     /isolate="TC3"
                     /isolation_source="clinical specimen"
                     /db_xref="taxon:239758"
                     /country="Taiwan"
     CDS             join(196. .13329,13329. .21416)
                     /note="ribosomal slippage"
                     /codon_start=1
                     /product="nonstructural polyprotein pplab"
                     /protein_id="AAP97879.1"
                     /db_xref="GI:33188325"
                     /translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
                     HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
                     TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEDYEQN
                     WNTKHGSGALRELTRELNGGAVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
                     LDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPFEIKSAKKFDTFKGECPKFVFP
                     LNSKVKVIQPRVEKKKTEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTC
                     DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSN
                     IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
                     LEILSRERVNINIVGDFHLNEEVAIILASFSASTSAFIDTIKSLDYKSFKTIVESCGN
                     YKVTKGKPVKGAWNIGQQRSVLTPLCGFPSQAAGVIRSIFARTLDAANHSIPDLQRAA
                     VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVQQTSQWLSNLLGTTVEKL
                     RPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKGQIQVASDNIKDCVKCFIDVV
                     NKALEMCIDQVTIAGAKLRSLNLGEVFIAQSKGLYRQCIRGKEQLQLLMPLKAPKEVT
                     FLEGDSHDTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTPVCVNGLMLLEIKDKEQ
                     YCALSPGLLATNNVFRLKGGAPIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNE
                     KCSVYTVESGTEVTEFACVVAEAVVKTLQPVSDLLTNMGIDLDEWSVATFYLFDDAGE
                     ENFSSRMYCSFYPPDEEEEDDAECEEEEIDETCEHEYGTEDDYQGLPLEFGASAETVR
                     VEEEEEEDWLDDTTEQSEIEPEPEPTPEEPVNQFTGYLKLTDNVAIKCVDIVKEAQSA
                     NPMVIVNAANIHLKHGGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL
                     AKKCLHVVGPNLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQT
                     VRTQVYIAVNDKALYEQVVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPVDV
                     KPKIKACIDEVTTTLEETKFLTNKLLLFADINGKLYHDSQNMLRGEDMSFLEKDAPYM
                     VGDVITSGDITCVVIPSKKAGGTTEMLSRALKKVPVDEYITTYPGQGCAGYTLEEAKT
                     ALKKCKSAFYVLPSEAPNAKEEILGTVSWNLREMLAHAEETRKLMPICMDVRAIMATI
                     QRKYKGIKIQEGIVDYGVRFFFYTSKEPVASIITKLNSLNEPLVTMPIGYVTHGFNLE
                     EAARCMRSLKAPAVVSVSSPDAVTTYNGYLTSSSKTSEEHFVETVSLAGSYRDWSYSG
                     QRTELGVEFLKRGDKIVYHTLESPVEFHLDGEVLSLDKLKSLLSLREVKTIKVFTTVD
                     NTNLHTQLVDMSMTYGQQFGPTYLDGADVTKIKPHVNHEGKTFFVLPSDDTLRSEAFE
                     YYHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVLLALQQLEVKFN
                     APALQEAYYRARAGDAANFCALILAYSNKTVGELGDVRETMTHLLQHANLESAKRVLN
                     VVCKHCGQKTTTLTGVEAVMYMGTLSYDNLKTGVSIPCVCGRDATQYLVQQESSFVMM
                     SAPPAEYKLQQGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTD
                     VFYKETSYTTTIKPVSYKLDGVTYTEIEPKLDGYYKKDNAYYTEQPIDLVPTQPLPNA
                     SFDNFKLTCSNTKFADDLNQMTGFTKPASRELSVTFFPDLNGDVVAIDYRHYSASFKK
                     GAKLLHKPIVWHINQATTKTTFKPNTWCLRCLWSTKPVDTSNSFEVLAVEDTQGMDNL
                     ACESQQPTSEEVVENPTIQKEVIECDVKTTEVVGNVILKPSDEGVKVTQELGHEDLMA
                     AYVENTSITIKKPNELSLALGLKTIATHGIAAINSVPWSKILAYVKPFLGQAAITTSN
                     CAKRLAQRVFNNYMPYVFTLLFQLCTFTKSTNSRIRASLPTTIAKNSVKSVAKLCLDA
                     GINYVKSPKFSKLFTIAMWLLLLSICLGSLICVTAAFGVLLSNFGAPSYCNGVRELYL
                     NSSNVTTMDFCEGSFPCSICLSGLDSLDSYPALETIQVTISSYKLDLTILGLAAEWVL
                     AYMLFTKFFYLLGLSAIMQVFFGYFASHFISNSWLMWFIISIVQMAPVSAMVRMYIFF
                     ASFYYIWKSYVHIMDGCTSSTCMMCYKRNRATRVECTTIVNGMKRSFYVYANGGRGFC
                     KTHNWNCLNCDTFCTGSTFISDEVARDLSLQFKRPINPTDQSSYIVDSVAVKNGALHL
                     YFDKAGQKTYERHPLSHFVNLDNLRANNTKGSLPINVIVFDGKSKCDESASKSASVYY
                     SQLMCQPILLLDQALVSDVGDSTEVSVKMFDAYVDTFSATFSVPMEKLKALVATAHSE
                     LAKGVALDGVLSTFVSAARQGVVDTDVDTKDVIECLKLSHHSDLEVTGDSCNNFMLTY
                     NKVENMTPRDLGACIDCNARHINAQVAKSHNVSLIWNVKDYMSLSEQLRKQIRSAAKK
                     NNIPFRLTCATTRQVVNVITTKISLKGGKIVSTCFKLMLKATLLCVLAALVCYIVMPV
                     HTLSIHDGYTNEIIGYKAIQDGVTRDIISTDDCFANKHAGFDAWFSQRGGSYKNDKSC
                     PVVAAIITREIGFIVPGLPGTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEYSDFA
                     TSACVLAAECTIFKDAMGKPVPYCYDTNLLEGSISYSELRPDTRYVLMDGSIIQFPNT
                     YLEGSVRVVTTFDAEYCRHGTCERSEVGICLSTSGRWVLNNEHYRALSGVFCGVDAMN
                     LIANIFTPLVQPVGALDVSASVVAGGIIAILVTCAAYYFMKFRRVFGEYNHVVAANAL
                     LFLMSFTILCLVPAYSFLPGVYSVFYLYLTFYFTNDVSFLAHLQWFAMFSPIVPFWIT
                     AIYVFCISLKHCHWFFNNYLRKRVMFNGVTFSTFEEAALCTFLLNKEMYLKLRSETLL
                     PLTQYNRYLALYNKYKYFSGALDTTSYREAACCHLAKALNDFSNSGADVLYQPPQTSI
                     TSAVLQSGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDTVYCPRHVICTAEDMLNP
                     NYEDLLIRKSNHSFLVQAGNVQLRVIGHSMQNCLLRLKVDTSNPKTPKYKFVRIQPGQ
                     TFSVLACYNGSPSGVYQCAMRPNHTIKGSFLNGSCGSVGFNIDYDCVSFCYMHHMELP
                     TGVHAGTDLEGKFYGPFVDRQTAQAAGTDTTITLNVLAWLYAAVINGDRWFLNRFTTT
                     LNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCAALKELLQNGMNGRTILGS
                     TILEDEFTPFDVVRQCSGVTFQGKFKKIVKGTHHWMLLTFLTSLLILVQSTQWSLFFF
                     VYENAFLPFTLGIMAIAACAMLLVKHKHAFLCLFLLPSLATVAYFNMVYMPASWVMRI
                     MTWLELADTSLSGYRLKDCVMYASALVLLILMTARTVYDDAARRVWTLMNVITLVYKV
                     YYGNALDQAISMWALVISVTSNYSGVVTTIMFLARAIVFVCVEYYPLLFITGNTLQCI
                     MLVYCFLGYCCCCYFGLFCLLNRYFRLTLGVYDYLVSTQEFRYMNSQGLLPPKSSIDA
                     FKLNIKLLGIGGKPCIKVATVQSKMSDVKCTSVVLLSVLQQLRVESSSKLWAQCVQLH
                     NDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINRLCEEMLDNRATLQAIASEFSSLPS
                     YAAYATAQEAYEQAVANGDSEVVLKKLKKSLNVAKSEFDRDAAMQRKLEKMADQAMTQ
                     MYKQARSEDKRAKVTSAMQTMLFTMLRKLDNDALNNIINNARDGCVPLNIIPLTTAAK
                     LMVVVPDYGTYKNTCDGNTFTYASALWEIQQVVDADSKIVQLSEINMDNSPNLAWPLI
                     VTALRANSAVKLQNNELSPVALRQMSCAAGTTQTACTDDNALAYYNNSKGGRFVLALL
                     SDHQDLKWARFPKSDGTGTIYTELEPPCRFVTDTPKGPKVKYLYFIKGLNNLNRGMVL
                     GSLAATVRLQAGNATEVPANSTVLSFCAFAVDPAKAYKDYLASGGQPITNCVKMLCTH
                     TGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHPNPKGFCDLKGKYVQIPTTCAND
                     PVGFTLRNTVCTVCGMWKGYGCSCDQLREPLMQSADASTFLNRVCGVSAARLTPCGTG
                     TSTDVVYRAFDIYNEKVAGFAKFLKTNCCRFQEKDEEGNLLDSYFVVKRHTMSNYQHE
                     ETIYNLVKDCPAVAVHDFFKFRVDGDMVPHISRQRLTKYTMADLVYALRHFDEGNCDT
                     LKEILVTYNCCDDDYFNKKDWYDFVENPDILRVYANLGERVRQSLLKTVQFCDAMRDA
                     GIVGVLTLDNQDLNGNWYDFGDFVQVAPGCGVPIVDSYYSLLMPILTLTRALAAESHM
                     DADLAKPLIKWDLLKYDFTEERLCLFDRYFKYWDQTYHPNCINCLDDRCILHCANFNV
                     LFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVNLHSSRLSFKELLV
                     YAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFYDFAVSKGFFKE
                     GSSVELKHFFFAQDGNAAISDYDYYRYNLPTMCDIRQLLFVVEVVDKYFDCYDGGCIN
                     ANQVIVNNLDKSAGFPFNKWGKARLYYDSMSYEDQDALFAYTKRNVIPTITQMNLKYA
                     ISAKNRARTVAGVSICSTMTNRQFHQKLLKSIAATRGATVVIGTSKFYGGWHNMLKTV
                     YSDVETPHLMGWDYPKCDRAMPNMLRIMASLVLARKHNTCCNLSHRFYRLANECAQVL
                     SEMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQAVTANVNALLSTDGNKIADKYVR
                     NLQHRLYECLYRNRDVDHEFVDEFYAYLRKHFSMMILSDDAVVCYNSNYAAQGLVASI
                     KNFKAVLYYQNNVFMSEAKCWTETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRI
                     LGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADVFHLYLQYIRKLHDEL
                     TGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTVLQAVGACVLCNSQTSLRCGACIR
                     RPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCDVTDVTQLYLGGMSYYCKSHKPPI
                     SFPLCANGQVFGLYKNTCVGSDNVTDFNAIATCDWTNAGDYILANTCTERLKLFAAET
                     LKATEETFKLSYGIATVREVLSDRELHLSWEVGKPRPPLNRNYVFTGYRVTKNSKVQI
                     GEYTFEKGDYGDAVVYRGTTTYKLNVGDYFVLTSHTVMPLSAPTLVPQEHYVRITGLY
                     PTLNISDEFSSNVANYQKVGMQKYSTLQGPPGTGKSHFAIGLALYYPSARIVYTACSH
                     AAVDALCEKALKYLPIDKCSRIIPARARVECFDKFKVNSTLEQYVFCTVNALPETTAD
                     IVVFDEISMATNYDLSVVNARLRAKHYVYIGDPAQLPAPRTLLTKGTLEPEYFNSVCR
                     LMKTIGPDMFLGTCRRCPAEIVDTVSALVYDNKLKAHKDKSAQCFKMFYKGVITHDVS
                     SAINRPQIGVVREFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQTVDSSQGSEYDY
                     VIFTQTTETAHSCNVNRFNVAITRAKIGILCIMSDRDLYDKLQFTSLEIPRRNVATLQ
                     AENVTGLFKDCSKIITGLHPTQAPTHLSVDIKFKTEGLCVDIPGIPKDMTYRRLISMM
                     GFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEGCHATRDAVGTNLPLQLGFSTGVN
                     LVAVPTGYVDTENNTEFTRVNAKPPPGDQFKHLIPLMYKGLPWNVVRIKIVQMLSDTL
                     KGLSDRVVFVLWAHGFELTSMKYFVKIGPERTCCLCDKRATCFSTSSDTYACWNHSVG
                     FDYVYNPFMIDVQQWGFTGNLQSNHDQHCQVHGNAHVASCDAIMTRCLAVHECFVKRV
                     DWSVEYPIIGDELRVNSACRKVQHMVVKSALLADKFPVLHDIGNPKAIKCVPQAEVEW
                     KFYDAQPCSDKAYKIEELFYSYATHHDKFTDGVCLFWNCNVDRYPANAIVCRFDTRVL
                     SNLNLPGCDGGSLYVNKHAFHTPAFDKSAFTNLKQLPFFYYSDSPCESHGKQVVSDID
                     YVPLKSATCITRCNLGGAVCRHHANEYRQYLDAYNMMISAGFSLWIYKQFDTYNLWNT
                     FTRLQSLENVAYNVVNKGHFDGHAGEAPVSIINNAVYTKVDGIDVEIFENKTTLPVNV
                     AFELWAKRNIKPVPEIKILNNLGVDIAANTVIWDYKREAPAHVSTIGVCTMTDIAKKP
                     TESACSSLTVLFDGRVEGQVDLFRNARNGVLITEGSVKGLTPSKGPAQASVNGVTLIG
                     ESVKTQFNYFKKVDGIIQQLPETYFTQSRDLEDFKPRSQMETDFLELAMDEFIQRYKL
                     EGYAFEHIVYGDFSHGQLGGLHLMIGLAKRSQDSPLKLEDFIPMDSTVKNYFITDAQT
                     GSSKCVCSVIDLLLDDFVEIIKSQDLSVISKVVKVTIDYAEISFMLWCKDGHVETFYP
                     KLQASQAWQPGVAMPNLYKMQRMLLEKCDLQNYGENAVIPKGIMMNVAKYTQLCQYLN
                     TLTLAVPYNMRVIHFGAGSDKGVAPGTAVLRQWLPTGTLLVDSDLNDFVSDADSTLIG
                     DCATVHTANKWDLIISDMYDPRTKHVTKENDSKEGFFTYLCGFIKQKLALGGSIAVKI
                     TEHSWNADLYKLMGHFSWWTAFVTNVNASSSEAFLIGANYLGKPKEQIDGYTMHANYI
                     FWRNTNPIQLSSYSLFDMSKFPLKLRGTAVMSLKENQINDMIYSLLEKGRLIIRENNR
                     VVVSSDILVNN"
     CDS             196. .13344
                     /codon_start=1
                     /product="nonstructural polyprotein pp1a"
                     /protein_id="AAP97880.1"
                     /db_xref="GI:33188326"
                     /translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
```

```
                     HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
                     TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEDYEQN
                     WNTKHGSGALRELTRELNGGAVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
                     LDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPFEIKSAKKFDTFKGECPKFVFP
                     LNSKVKVIQPRVEKKKTEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTC
                     DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSN
                     IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
                     LEILSRERVNINIVGDFHLNEEVAIILASFSASTSAFIDTIKSLDYKSFKTIVESCGN
                     YKVTKGKPVKGAWNIGQQRSVLTPLCGFPSQAAGVIRSIFARTLDAANHSIPDLQRAA
                     VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVQQTSQWLSNLLGTTVEKL
                     RPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKGQIQVASDNIKDCVKCFIDVV
                     NKALEMCIDQVTIAGAKLRSLNLGEVFIAQSKGLYRQCIRGKEQLQLLMPLKAPKEVT
                     FLEGDSHDTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTPVCVNGLMLLEIKDKEQ
                     YCALSPGLLATNNVFRLKGGAPIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNE
                     KCSVYTVESGTEVTEFACVVAEAVVKTLQPVSDLLTNMGIDLDEWSVATFYLFDDAGE
                     ENFSSRMYCSFYPPDEEEEDDAECEEEEIDETCEHEYGTEDDYQGLPLEFGASAETVR
                     VEEEEEEDWLDDTTEQSEIEPEPEPTPEEPVNQFTGYLKLTDNVAIKCVDIVKEAQSA
                     NPMVIVNAANIHLKHGGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL
                     AKKCLHVVGPNLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQT
                     VRTQVYIAVNDKALYEQVVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPVDV
                     KPKIKACIDEVTTTLEETKFLTNKLLLFADINGKLYHDSQNMLRGEDMSFLEKDAPYM
                     VGDVITSGDITCVVIPSKKAGGTTEMLSRALKKVPVDEYITTYPGQGCAGYTLEEAKT
                     ALKKCKSAFYVLPSEAPNAKEEILGTVSWNLREMLAHAEETRKLMPICMDVRAIMATI
                     QRKYKGIKIQEGIVDYGVRFFFYTSKEPVASIITKLNSLNEPLVTMPIGYVTHGFNLE
                     EAARCMRSLKAPAVVSVSSPDAVTTYNGYLTSSSKTSEEHFVETVSLAGSYRDWSYSG
                     QRTELGVEFLKRGDKIVYHTLESPVEFHLDGEVLSLDKLKSLLSLREVKTIKVFTTVD
                     NTNLHTQLVDMSMTYGQQFGPTYLDGADVTKIKPHVNHEGKTFFVLPSDDTLRSEAFE
                     YYHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVLLALQQLEVKFN
                     APALQEAYYRARAGDAANFCALILAYSNKTVGELGDVRETMTHLLQHANLESAKRVLN
                     VVCKHCGQKTTTLTGVEAVMYMGTLSYDNLKTGVSIPCVCGRDATQYLVQQESSFVMM
                     SAPPAEYKLQQGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTD
                     VFYKETSYTTTIKPVSYKLDGVTYTEIEPKLDGYYKKDNAYYTEQPIDLVPTQPLPNA
                     SFDNFKLTCSNTKFADDLNQMTGFTKPASRELSVTFFPDLNGDVVAIDYRHYSASFKK
                     GAKLLHKPIVWHINQATTKTTFKPNTWCLRCLWSTKPVDTSNSFEVLAVEDTQGMDNL
                     ACESQQPTSEEVVENPTIQKEVIECDVKTTEVVGNVILKPSDEGVKVTQELGHEDLMA
                     AYVENTSITIKKPNELSLALGLKTIATHGIAAINSVPWSKILAYVKPFLGQAAITTSN

  Query Match             90.9%;  Score 20;  DB 14;  Length 29573;
  Best Local Similarity   75.0%;  Pred. No. 6.5;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        707 GTGAACTCACTCGTGAGCTC 726


RESULT 14
AY559094
LOCUS       AY559094               29577 bp    RNA     linear   VRL 24-SEP-2004
DEFINITION  SARS coronavirus Sin846, complete genome.
ACCESSION   AY559094
VERSION     AY559094.1  GI:45645021
KEYWORDS    .
SOURCE      SARS coronavirus Sin846
  ORGANISM  SARS coronavirus Sin846
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29577)
  AUTHORS   Vega,V.B., Ruan,Y., Liu,J., Lee,W.H., Wei,C.L., Se-Thoe,S.Y.,
            Tang,K.F., Zhang,T., Kolatkar,P.R., Ooi,E.E., Ling,A.E.,
            Stanton,L.W., Long,P.M. and Liu,E.T.
  TITLE     Mutational dynamics of the SARS coronavirus in cell culture and
            human populations isolated in 2003
  JOURNAL   (er) BMC Infect. Dis. 4 (1), 32 (2004)
   PUBMED   15347429
REFERENCE   2  (bases 1 to 29577)
  AUTHORS   Vega,V.B., Ruan,Y., Liu,J., Lee,W.H., Wei,C.L., Se Thoe,S.Y.,
            Tang,K.F., Zhang,T., Lin,S., Kolatkor,P., Ooi,E.E., Ee,L.A.,
            Stanton,L.W., Long,P.M. and Liu,E.T.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-2004) Genome Institute of Singapore, 60, Biopolis
            Street #02-01, Genome, Singapore 1386782, Singapore
FEATURES             Location/Qualifiers
     source          1. .29577
                     /organism="SARS coronavirus Sin846"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:267396"
                     /country="Singapore"
ORIGIN

  Query Match             90.9%;  Score 20;  DB 14;  Length 29577;
  Best Local Similarity   75.0%;  Pred. No. 6.5;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        760 GTGAACTCACTCGTGAGCTC 779


RESULT 15
AY463059
LOCUS       AY463059               29592 bp    RNA     linear   VRL 05-JAN-2004
DEFINITION  SARS coronavirus ShanghaiQXC1, complete genome.
ACCESSION   AY463059 AH013000 AY322197 AY322198 AY322199
VERSION     AY463059.1  GI:40457433
KEYWORDS    .
SOURCE      SARS coronavirus ShanghaiQXC1
  ORGANISM  SARS coronavirus ShanghaiQXC1
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29592)
  AUTHORS   Yuan,Z., Zhang,X., Hu,Y., Lan,S., Zhou,Z., Wang,H. and Wen,Y.
  TITLE     Analysis of SARS coronavirus genome in Shanghai isolates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 29592)
  AUTHORS   Yuan,Z., Zhang,X., Hu,Y., Lan,S., Zhou,Z., Wang,H. and Wen,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-2003) Key Lab of Medical Molecular Virology,
            Shanghai Medical College, Fudan University, 138 Yi Xue Yuan Road,
            Shanghai 200032, P.R. China
COMMENT     On or before Jan 5, 2004 this sequence version replaced
            gi:32454355, gi:32454356, gi:32454357, gi:32454358.
FEATURES             Location/Qualifiers
     source          1. .29592
                     /organism="SARS coronavirus ShanghaiQXC1"
                     /virion
                     /mol_type="genomic RNA"
                     /strain="ShanghaiQXC1"
                     /db_xref="taxon:258507"
     CDS             join(186. .13313,13313. .21406)
                     /note="ribosomal slippage"
                     /codon_start=1
                     /product="orf1ab"
                     /protein_id="AAP82978.2"
                     /db_xref="GI:40557708"
                     /translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
                     HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
                     TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEDYEQN
                     WNTKHGSGALRELTRELNGGAVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
                     LDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPFEIKSAKKFDTFKGECPKFVFP
                     LNSKVKVIQPRVEKKKTEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTC
                     DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSN
                     IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
                     LEILSRERVNINIVGDFHLNEEVAIILASFSASTSAFIDTIKSLDYKSFXTIVESCGN
                     YKVTKGKPVKGAWNIGQQRSVLTPLCGFPSQAAGVIRSIFARTLDAANHSIPDLQRAA
                     VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVQQTSQWLSNLLGTTVEKL
                     RPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKGQIQVASDNIKDCVKCFIDVV
                     NKALEMCIDQVTIAGAKLRSLNLGEVFIAQSKGLYRQCIRGKEQLQLLMPLKAPKEVT
                     FLEGDSHDTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTPVCVNGLMLLEIKDKEQ
                     YCALSPGLLATNNVFRLKGGAPIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNE
                     KCSVYTVESGTEVTEFACVVAEAVVKTLQPVSDLLTNMGIDLDEWSVATFYLFDDAGE
                     ENFSSRMYCSFYPPDEEEEDDAECEEEEIDETCEHEYGTEDDYQGLPLEFGASAETVR
                     VEEEEEEDWLDDTTEQSEIEPEPEPTPEEPVNQFTGYLKLTDNVAIKCVDTVKEAQSA
                     NPMVIVNAANIHLKHGGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL
                     AKKCLHVVGPNLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQT
                     VRTQVYIAVNDKALYEQVVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPVDV
```

KPKIKACIDEVTTTLEETKFLTNKLLLFADINGKLYHDSQNMLRGEDMSFLEEDAPYM
VGDVITSGDITCVVIPSKKAGGTTEMLSRALKKVPVDEYITTYPGQGCAGYTLEEAKT
ALKKCKSAFYVLPSEAPNAKEEILGTVSWNLREMLAHAEEARKLMPICMDVRAIMATI
QRKYKGVKIQEGIVDYGVRFFFYTSKEPVASIITKLNSLNEPLVTMPIGYVTHGFNLE
EAARCMRSLKAPAVVSVSSPDAVTTYNGYLTSSSKTSEEHFVETVSLAGSYRDWSYSG
QRTELGVEFLKRGDKIVYHTLESPVEFHLDGEVLSLDKLKSLLSLREVKTIKVFTTVD
NTNLHTQLVDMSMTYGQQFGPTYLDGADVTKIKPHVNHEGKTFFVLPSDDTLRSEAFE
YYHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVLLALQQLEVKFN
APALQEAYYRARAGDAANFCALILAYSNKTVGELGDVRETMTHLLQHANLESAKRVLN
VVCKHCGQKTTTLTGVEAVMYMGTLSYDNLKTGVSIPCVCGRDATQYLVQQESSFVMM
SAPPAGYKLQQGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTD
VFYKETSYTTTIKPVSYKLDGVTYTEIEPKLDGYYKKDNAYYTEQPIDLVPTQPLPNA
SFDNFKLTCSNTKFADDLNQMTGFTKPASRELSVTFFPDLNGDVVAIDYRHYSASFKK
GAKLLHKPIVWHINQATTKTTFKPNTWCLRCLWSTKPVDTSNSFEVLAVEDTQGMDNL
ACESQQPTSEEVVENPTIQKEVIECDVKTTEVVGNVILKPSDEGVKVTQELGHEDLMA
AYVENTSITIKKPNELSLALGLKTIATHGIAAINSVPWSKILAYVKPFLGQAAITTSN
CAKRLAQRVFNNYMPYVFTLLFQLCTFTKSTNSRIRASLPTTIAKNSVKSVAKLCLDA
GINYVKSPKFSKLFTIAMWLLLLSICLGSLICVTAAFGVLLSNFGAPSYCNGVRELYL
NSSNVTTMDFCEGSFPCSICLSGLDSLDSYPALETIQVTISSYKLDLTILGLAAEWVL
AYMLFTKFFYLLGLSAIMQAFFGYFASHFISNSWLMWFIISIVQMAPVSAMVRMYIFF
ASFYYIWKSYVHIMDGCTSSTCMMCYKCNHATRVECTTIVNGMKRSFYVYANGGRGFC
KTHNWNCLNCDTFCTGSTFISDEVARDLSLQFKRPINPTDQSSYIVDSVAVKNGALHP
YFDKAGQKTYERHPLSHFVNLDNLRANNTKGSLPINVIVFDGKSKCDESASKSASVYY
SQLMCQPILLLDQALVSDVGDSTEVSVKMFDAYVDTFSATFSVPMEKLKALVATAHSE
LAKGVALDGVLSTFVSAARQGVVDTDVDTKDVIECLKLSHHSDLEVTGDSCYNFMLTY
NKVENMTPRDLGACIDCNARHINAQVAKSHNVSLIWNVKDYMSLSEQLRKQIRSAAKK
NNIPFRLTCATTRQVVNVITTKISLKGGKIVSTCFKLMLKATLLCVLAALVCYIVMPV
HTLSIHDGYTNEIIGYKAIQDGVTRDIISTDDCFANKHAGFDAWFSQRGGSYKNDKSC
PVVAAIITREIGFIVPGLPGTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEYSDFA
TSACVLAAECTIFKDAMGKPVPYCYDTNLLEGSISYSELRPDTRYVLMDGSIIQFPNT
YLEGSVRVVTTFDAEYCRHGTCERSEVGICLSTSGRWVLNNEHYRALSGVFCGVDAMN
LIANIFTPLVQPVGALDASASVVAGGIIAILVTCAAYYFMKFRRAFGEYNHVVAANAL
LFLMSFTILCLVPAYSFLPGVYSVFYLYLTFYFTNDVSFLAHLQWFAMFSPIVPFWIT
AIYVFCISLKHCHWFFNNYLRKRVMFNGVTFSTFEEAALCTFLLNKEMYLKLRSETLL
PLTQYNRYLALYNKYKYFSGVLDTTSYREAACCHLAKALNDFSNSGADVLYQPPQTSI
TSAVLQSGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDTVYCPRHVICTAEDMLNP
NYEDLLIRKSNHSFLVQAGNVQLRVIGHSMQNCLLRLKVDTSNPKTPKYKFVRIQPGQ
TFSVLACYNGSPSGVYQCAMRPNHTIKGSFLNGSCGSVGFNIDYDCVSFCYMHHMELP
TGVHAGTDLEGKFYGPFVDRQTAQAAGTDTTITLNVLAWLYAAVINGDRWFLNRFTTT
LNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCAALKELLQNGMNGRTILGS
TILEDEFTPFDVVRQCSGVTFQGKFKKIVKGTHHWMLLTFLTSLLILVQSTQWSLFFF
VYENAFLPFTLGIMAIAACAMLLVKHKHAFLCLFLLPSLATVAYFNMVYMPASWVMRI
MTWLELADTSLSGYRLKDCVMYASALVLLILMTARTVYDDAARRVWTLMNVITLVYKV
YYGNALDQAISMWALVISATSNYSGVVTTIMFLARAIVFVCVEYYPLLFITGNTLQCI
MLVYCFLGYCCCCYFGLFCLLNRYFRLTLGVYDYLVSTQEFRYMNSQGLLPPKSSIDA
FKLNIKLLGIGGKPCIKVATVQSKMSDVKCTSVVLLSVLQQLRVESSSKLWAQCVQLH
NDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINRLCEEVLDNRATLQAIASEFSSLPS
YAAYATAQEAYEQAVANGDSEVVLKKLKKSLNVAKSEFDRDAAMQRKLEKMADQAMTQ
MYKQARSEDKRAKVTSAMQTMLFTMLRKLDNDALNNIINNARDGCVPLNIIPLTTAAK
LMVVVPDYGTYKNTCDGNTFTYASALWEIQQVVDADSKIVQLSEINMDNSPNLAWPLI
VTALRANSAAKLQNNELSPVALRQMSCAAGTTQTACTDDNALAYYNNSKGGRFVLALL
SDHQDLKWARFPKSDGTGTIYTELEPPCRFVTDTPKGPKMKYLYFIKGLNNLNRGMVL
GSLAATVRLQAGNATEVPANSTVLSFCAFAVDPAKAYKDYLASGGQPITNCVKMLCTH
TGTGQAITVTPEANMDQGSFGGASCCLYCRCHIDHPNPKGFCDLKGKYVQIPTTCAND
PVGFTLRNTVCTVCGMWKGYGCSCDQLREPLMQSADASTFFKRVCGVSAARLTPCGTG
TSTDVVYRAFDIYNEKVAGFAKFLKTNCCRFQEKDEEGNLLDSYFVVKRHTMSNYQHE
ETIYNLIKDCPAVAVHDFFKFRVDGDMVPHISRQRLTKYTMADLVYALRHFDEGNCDT
LKEILVTYNCCDDDYFNKKDWYDFVENPDILRVYANLGERVRQSLLKTVQFCDAMRDA
GIVGVLTLDNQDLNGNWYDFGDFVQVAPGCGVPIVDSYYSLLMPILTLTRALAAESHM
DADLAKPLIKWDLLKYDFTEERLCLFDRYFKYWDQTYHPNCINCLDDRCILHCANFNV
LFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVNLHSSRLSFKELLV
YAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFYDFAVSKGFFKE
GSSVELKHFFFAQDGNAAISDYDYYRYNLPTMCDIRQLLFVVEVVDKYFDCYDGGCIN
ANQVIVNNLDKSAGFPFNKWGKARLYYDSMSYEDQDALFAYTKRNVIPTITQMNLKYA
ISAKNRARTVAGVSICSTMTNRQFHQKLLKSIAATRGATVVIGTSKFYGGWHNMLKTV
YSDVETPHLMGWDYPKCDRAMPNMLRIMASLVLARKHNTCCNLSHRFYRLANECAQVL
SEMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQAVTANVNALLSTDGNKIADKYVR
NLQHRLYECLYRNRDVDHEFVDEFYAYLRKHFSMMILSDDAVVCYNSNYAAQGLVASI
KNFKAVLYYQNNVFMSEAKCWTETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRI
LGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADVFHLYLQYIRKLHDEL
TGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTVLQAVGACVLCNSQTSLRCGACIR
RPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCDVTDVTQLYLGGMSYYCKSHKPPI
SFPLCANGQVFGLYKNTCVGSDNVTDFNAIATCDWTNAGDYILANTCTERLKLFAAET

LKATEETFKLSYGIATVREVLSDRELHLSWEVGKPRPPLNRNYVFTGYRVTKNSKVQI
GEYTFEKGDYGDAVVYRGTTTYKLNVGDYFVLTSHTVMPLSAPTLVPQEHYVRITGLY
PTLNISDEFSSNVANYQKVGMQKYSTLQGPPGTGKSHFAIGLALYYPSARIVYTACSH
AAVDALCEKALKYLPIDKCSRIIPARARVECFDKFKVNSTLEQYVFCTVNALPETTAD
IVVFDEISMATNYDLSVVNARLRAKHYVYIGDPAQLPARTLLTKGTLEPEYFNSVCR
LMKTIGPDMFLGTCRRCPAEIVDTVSALVYDNKLKAHKEKSAQCFKMFYKGVITHDVS
SAINRPQIGVVREFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQTVDSSQGSEYDY
VIFTQTTETAHSCNVNRFNVAITRAKIGILCIMSDRDLYDKLQFTSLEIPRRNVATLQ
AENVTGLFKDCSKIITGLHPTQAPTHLSVDIKFKTEGLCVDIPGIPKDMTYRRLISMM
GFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEGCHATRDAVGTNLPLQLGFSTGVN
LVAVPTGYVDTENNTEFTRVNAKPPPGDQFKHLIPLMYKGLPWNVVRIKIVQMLSDTL
KGLSDRVVFVLWAHGFELTSMKYFVKIGPERTCCLCDKRATCFSTSSDTYACWNHSVG
FDYVYNPFMIDVQQWGFTGNLQSNHDQHCQVHGNAHVASCDAIMTRCLAVHECFVKRV
DWSVEYPIIGDELRVNSACRKVQHMVVKSALLADKFPVLHDIGNPKAIKCVPQAEVEW
KFYDAQPCSDKAYKIEELFYSYATHHDKFTDGVCLFWNCNVDRYPANAIVCRFDTRVL
SNLNLPGCDGGSLYVNKHAFHTPAFDKSAFTNLKQLPFFYYSDSPCESHGKQVVSDID
YVPLKSATCITRCNLGGAVCRHHANEYRQYLDAYNMMISAGFSLWIYKQFDTYNLWNT
FTRLQSLENVAYNVVNKGHFDGHAGEAPVSIINNAVYTKVDGIDVEIFESKTTLPVNV
AFELWAKRNIKPVPEIKILNNLGVDIAANTVIWDYKREAPAHVSTIGVCTMTDIAKKP
TESACSSLTVLFDGRVEGQVDLFRNARNGVLITEGSVKGLTPSKGPAQASVNGVTLIG
ESVKTQFNYFKKVDGIIQQLPETYFTQSRDLEDFKPRSQMETDFLELAMDEFIQRYKL
EGYAFEHIVYGDFSHGQLGGLHLMIGLAKRSQDSPLKLEDFIPMDSTVKNYFITDAQT
GSSKCVRSVIDLLLDDFVEIIKSQDLSVISKVVKVTIDYAEISFMLWCKDGHVETFYP
KLQASQAWQPGVAMPNLYKMQRMLLEKCDLQNYGENAVIPKGIMMNVAKYTQLCQYLN
TLTLAVPYNMRVIHFGAGSDKGVAPGTAVLRQWLPTGTLLVDSDLNDFVSDADSTLIG
DCATVHTANKWDLIISDMYDPRTKHVTKENDSKEGFFTYLCGFIKQKLALGGSIAVKI
TEHSWNADLYKLMGHFSWWTAFVTNVNASSSEAFLIGANYLGKPKEQIDGYTMHANYI
FWRNTNPIQLSSYSLFDMSKFPLKLRGTAVMSLKENQINDMIYSLLEKGRLIIRENNR
VVVSSDILVNN"
CDS 21413. .25180
/codon_start=1
/product="spike"
/protein_id="AAR86788.1"
/db_xref="GI:40457434"
/translation="MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPD
EIFRSDTLYLTQDLFLPFYSNVTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRGWV
FGSTMNNKSQSVIIINNSTNVVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCT
FEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLKP
IFKLPLGINITNFRAILTAFSPAQDIWGTSAAAYFVGYLKPTTFMLKYDENGTITDAV
DCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKF
PSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGD
DVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRP
FERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPA
TVCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPK
TSEILDISPCAFGGVSVITPGTNASSEVAVLYQDVNCTNVSAAIHADQLTPAWRIYST
GNNVFQTQAGCLIGAEHVDTSYECDIPIGAGICASYHTVSLLRSTSQKSIVAYTMSLG
ADSSIAYSNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGS
FCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRS
FIEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYT
AALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAIS
QIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAE
VQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYH
LMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQ
RNFFSPQIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKGELDKYFKNHTSPDV
DLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIA
GLIAIVMVTILLCCMTSCCSCLKGACSCGSCCKFDEDDSEPVLKGVKLHYT"
CDS 25189. .26013
/codon_start=1
/product="unknown"
/protein_id="AAP82984.2"
/db_xref="GI:40557709"
/translation="MDLFMRFFTLGSITAQPVKIDNASPASTAHATATIPLQASLPFG
WLVIGVAFLAVFQSATKIIALNKRWQLALYKGFQFICNLLLLFVTIYSHLLLVAAGME
AQFLYLYALIYFLQCINACRIIMRCWLCWKCKSKNPLLYDANYFVCWHTHNYDYCIPY
NSVTDTIVVTEGDGISTPKLKEDYQIGGYSEDRHSGVKDYVVVHGYFTEVYYQLESTQ
ITTNTGIENATFFIFNKLVKDPPNVQIHTIDGSSGVANPAMDPIYDEPTTTTSVPL"
CDS 25610. .26074
/codon_start=1
/product="unknown"
/protein_id="AAR86789.1"
/db_xref="GI:40457435"
/translation="MMPTTLFAGTHITMTTVYHITVSQIQLSLLKVTAFQHQNSKKTT

Query Match 90.9%; Score 20; DB 14; Length 29592;
Best Local Similarity 75.0%; Pred. No. 6.5;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        697 GTGAACTCACTCGTGAGCTC 716
```

Search completed: June 24, 2005, 01:35:35
Job time : 765.417 secs

THIS PAGE IS BLANK

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 23, 2005, 23:06:41 ; Search time 192.283 Seconds
(without alignments)
677.304 Million cell updates/sec

Title: US-10-848-737-1
Perfect score: 22
Sequence: 1 gugaacucacucgugagcuctt 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		20	90.9	29751	12	ADJ39000	Adj39000 SARS coro
2		18.8	85.5	340449	8	AAL52198	Aal52198 Human sec
3		17.2	78.2	4113	10	ABT41930	Abt41930 Toxicity
4		17.2	78.2	4113	12	ADP72828	Adp72828 Renal tox
5		17.2	78.2	4131	2	AAV26965	Aav26965 Rat kidne
6		17.2	78.2	4131	2	AAT95860	Aat95860 Rat kidne
7		17.2	78.2	4131	2	AAZ25056	Aaz25056 Rat parat
8		17.2	78.2	4131	2	AAV82486	Aav82486 Rat parat
9		17.2	78.2	4131	3	AAZ89299	Aaz89299 Rat calci
10		17.2	78.2	4131	6	AAI72123	Aai72123 cDNA enco
11		17.2	78.2	4550	12	ADO30115	Ado30115 Mouse GPC
12		17.2	78.2	57561	11	ACN44600	Acn44600 Mouse gen
	c 13	17.2	78.2	110000	13	ABD32923_5	Continuation (6 of
14		16.8	76.4	44	2	AAQ21468	Aaq21468 ZC2938- P
15		16.8	76.4	44	2	AAQ59009	Aaq59009 Primer to
16		16.8	76.4	44	2	AAV70902	Aav70902 Primer ZC
17		16.8	76.4	305	6	ABV88798	Abv88798 Human col
	c 18	16.8	76.4	398	8	ABX55821	Abx55821 Bovine ES
19		16.8	76.4	534	6	ABN69216	Abn69216 Streptoco
20		16.8	76.4	537	8	ACA50276	Aca50276 Prokaryot
	c 21	16.8	76.4	539	4	AAK91848	Aak91848 Human cDN
	c 22	16.8	76.4	539	4	AAK93253	Aak93253 Human cDN
	c 23	16.8	76.4	539	12	ADL28275	Adl28275 5' end of
	c 24	16.8	76.4	539	12	ADL29680	Adl29680 5' end of
	c 25	16.8	76.4	945	13	ADT44529	Adt44529 Bacterial
	c 26	16.8	76.4	1065	5	AAS70168	Aas70168 DNA encod
	c 27	16.8	76.4	1556	4	AAI63946	Aai63946 Human pol
	c 28	16.8	76.4	1556	4	AAS31613	Aas31613 cDNA enco
	c 29	16.8	76.4	1556	4	ABK43947	Abk43947 DNA encod
	c 30	16.8	76.4	1556	12	ADI54334	Adi54334 cDNA enco
	c 31	16.8	76.4	1556	12	ADM24497	Adm24497 Human PRO
	c 32	16.8	76.4	2518	4	AAK94748	Aak94748 Human ful
	c 33	16.8	76.4	2518	12	ADL31785	Adl31785 Full leng
	c 34	16.8	76.4	2991	4	ABK43659	Abk43659 DNA encod
	c 35	16.8	76.4	2991	12	ADI54046	Adi54046 cDNA enco
	c 36	16.8	76.4	3849	4	AAI63871	Aai63871 Human pol
	c 37	16.8	76.4	3849	12	ADM24422	Adm24422 Human PRO
	c 38	16.8	76.4	5195	10	ADF59868	Adf59868 Human con
39		16.8	76.4	12505	10	ADE62909	Ade62909 Rat gene
	c 40	16.8	76.4	95484	12	ADQ97298	Adq97298 Mouse can
41		16.4	74.5	350	8	ABZ56738	Abz56738 Aspergill
42		16.4	74.5	619	10	ADK58384	Adk58384 Plant DNA
43		16.4	74.5	619	10	ADK57664	Adk57664 Plant DNA
44		16.4	74.5	619	11	ADM45449	Adm45449 Insect re
45		16.4	74.5	633	3	AAF13945	Aaf13945 Aspergill

ALIGNMENTS

RESULT 1
ADJ39000
ID ADJ39000 standard; DNA; 29751 BP.
XX
AC ADJ39000;
XX
DT 06-MAY-2004 (first entry)
XX
DE SARS coronavirus nucleotide sequence.
XX
KW small interfering RNA; siRNA; modified ribonucleotide;
KW viral replication inhibition; hepatitis C virus; HCV; hepatitis C;
KW antiinflammatory; hepatotropic; virucide; hepatitis A virus;
KW hepatitis D virus; hepatitis E virus; Ebola virus; influenza virus;
KW rotavirus; reovirus; retrovirus; poliovirus; human papilloma virus;
KW metapneumoniavirus; coronavirus; viral infection; gene; ds.
XX
OS SARS coronavirus.
XX
PN WO2004011647-A1.
XX
PD 05-FEB-2004.
XX
PF 25-JUL-2003; 2003WO-US023104.
XX
PR 26-JUL-2002; 2002US-0398605P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Han J, Seo MY, Houghton M;
XX
DR WPI; 2004-143862/14.
XX
PT New RNase resistant small interfering RNA, useful for treating viral
PT infections, e.g., hepatitis C, influenza virus or coronavirus infection.
XX
PS Example 10; Fig 3; 74pp; English.
XX
CC The present invention describes a small interfering RNA (siRNA) which
CC comprises a modified ribonucleotide, where the siRNA is resistant to
CC RNase and retains the ability to inhibit viral replication. Also
CC described: (1) inactivating a virus in a patient; (2) making a modified
CC siRNA that targets a nucleic acid sequence in a virus; (3) a double-

```
CC   stranded RNA molecule of 10-30 nucleotides that inhibits replication of
CC   hepatitis C virus (HCV); (4) inducing targeted RNA interference toward
CC   HCV in hepatic cells; (5) inhibiting replication of HCV; (6) a vector
CC   comprising a DNA segment encoding the RNA molecule; (7) a host cell
CC   comprising the vector of (6); (8) inhibiting replication of HCV in cells
CC   carrying HCV; (9) treating hepatitis C in a subject; (10) a modified
CC   siRNA molecule comprising a double-stranded RNA molecule of 10-30
CC   nucleotides in length, which mediates RNA interference toward a target
CC   agent or virus and is linked to at least one receptor-binding ligand; and
CC   (11) inducing targeted RNA interference in a patient. The modified siRNA
CC   molecules have antiinflammatory, hepatotropic and virucide activities.
CC   The modified RNA molecules are useful for inactivating virus in mammalian
CC   cells. The siRNAs are useful for treating hepatitis C virus, hepatitis A
CC   virus, hepatitis D virus, hepatitis E virus, Ebola virus, influenza
CC   virus, rotavirus, reovirus, retrovirus, poliovirus, human papilloma
CC   virus, metapneumoniavirus or coronavirus infections. The methods of the
CC   invention can be used to correct or compensate for cellular physiological
CC   abnormalities involved in conferring susceptibility to viral infections
CC   in patients and/or alleviate symptoms of a viral infection in patients.
CC   The present sequence represents the SARS coronavirus nucleotide sequence,
CC   which is used in an example from the present invention.
XX
SQ   Sequence 29751 BP; 8478 A; 5941 C; 6188 G; 9144 T; 0 U; 0 Other;

  Query Match             90.9%;  Score 20;  DB 12;  Length 29751;
  Best Local Similarity   75.0%;  Pred. No. 8;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        776 GTGAACTCACTCGTGAGCTC 795


RESULT 2
AAL52198
ID   AAL52198 standard; cDNA; 340449 BP.
XX
AC   AAL52198;
XX
DT   22-SEP-2003  (first entry)
XX
DE   Human secreted protein genomic DNA coding sequence.
XX
KW   Human; gene; ds; secreted protein; chromosome 5; tissue typing;
KW   secreted protein-related disease; transgenic animal; drug screening;
KW   pharmacogenomic analysis; single nucleotide polymorphism; SNP.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(1107, T)
FT                   /*tag= a
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(5479, G)
FT                   /*tag= b
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(5480, G)
FT                   /*tag= c
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(6100, T)
FT                   /*tag= d
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(6122, A)
FT                   /*tag= e
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(6883, G)
FT                   /*tag= f
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(7641, G)
FT                   /*tag= g
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(8493, T)
FT                   /*tag= h
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(9626, A)
FT                   /*tag= i
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(9796, A)
FT                   /*tag= j
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(9868, A)
FT                   /*tag= k
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(10209, C)
FT                   /*tag= l
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(10311, G)
FT                   /*tag= m
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(10831, T)
FT                   /*tag= n
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(12230, C)
FT                   /*tag= o
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(13378, C)
FT                   /*tag= p
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(13423, T)
FT                   /*tag= q
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(13425, G)
FT                   /*tag= r
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(14116, G)
FT                   /*tag= s
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(14508, C)
FT                   /*tag= t
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(14645, A)
FT                   /*tag= u
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(14932, A)
FT                   /*tag= v
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(16272, G)
FT                   /*tag= w
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(16945, T)
FT                   /*tag= x
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(17065, G)
FT                   /*tag= y
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(17442, G)
FT                   /*tag= z
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(17475, C)
FT                   /*tag= aa
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(17844, A)
FT                   /*tag= ab
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(17929, G)
FT                   /*tag= ac
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(18540, C)
FT                   /*tag= ad
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(18592, T)
FT                   /*tag= ae
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(18616, G)
FT                   /*tag= af
FT                   /standard_name= "Single nucleotide polymorphism"
```

```
FT   variation       replace(20491, T)
FT                   /*tag= ag
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(22023, G)
FT                   /*tag= ah
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(22248, G)
FT                   /*tag= ai
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(23118, C)
FT                   /*tag= aj
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(23789, A)
FT                   /*tag= ak
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(23800, A)
FT                   /*tag= al
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(23930, A)
FT                   /*tag= am
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(24159, G)
FT                   /*tag= an
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(26339, A)
FT                   /*tag= ao
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(26508, G)
FT                   /*tag= ap
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(26570, A)
FT                   /*tag= aq
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(26582, G)
FT                   /*tag= ar
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(26693, A)
FT                   /*tag= as
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(26884, T)
FT                   /*tag= at
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(27320, T)
FT                   /*tag= au
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(27339, A)
FT                   /*tag= av
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(27542, G)
FT                   /*tag= aw
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(28586, A)
FT                   /*tag= ax
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(28591, T)
FT                   /*tag= ay
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(28599, A)
FT                   /*tag= az
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(30857, A)
FT                   /*tag= ba
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(31503, T)
FT                   /*tag= bb
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(33671, A)
FT                   /*tag= bc
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(35045, C)
FT                   /*tag= bd
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(35944, G)
FT                   /*tag= be
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(37157, G)
FT                   /*tag= bf
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(37603, T)
FT                   /*tag= bg
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(39242, G)
FT                   /*tag= bh
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(39404, A)
FT                   /*tag= bi
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(40395, C)
FT                   /*tag= bj

  Query Match             85.5%;  Score 18.8;  DB 8;  Length 340449;
  Best Local Similarity   68.2%;  Pred. No. 48;
  Matches   15;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||:| |:| :||||:|||
Db     267426 GTGAACTCTCTCCTGAGCTCTT 267447


RESULT 3
ABT41930
ID   ABT41930 standard; DNA; 4113 BP.
XX
AC   ABT41930;
XX
DT   26-JUN-2003  (first entry)
XX
DE   Toxicity modelling related rat gene SEQ ID No 1632.
XX
KW   Toxic effect; gene expression profile; renal toxicity; toxicity marker;
KW   database; drug screening; toxicity assay; rat; ds.
XX
OS   Rattus norvegicus.
XX
PN   WO200295000-A2.
XX
PD   28-NOV-2002.
XX
PF   22-MAY-2002; 2002WO-US016173.
XX
PR   22-MAY-2001; 2001US-0292335P.
PR   13-JUN-2001; 2001US-0297523P.
PR   19-JUN-2001; 2001US-0298925P.
PR   10-JUL-2001; 2001US-0303807P.
PR   10-JUL-2001; 2001US-0303808P.
PR   10-JUL-2001; 2001US-0303810P.
PR   28-AUG-2001; 2001US-0315047P.
PR   27-SEP-2001; 2001US-0324928P.
PR   22-OCT-2001; 2001US-0330462P.
PR   01-NOV-2001; 2001US-0330867P.
PR   21-NOV-2001; 2001US-0331805P.
PR   06-DEC-2001; 2001US-0336144P.
PR   19-DEC-2001; 2001US-0340873P.
PR   21-FEB-2002; 2002US-0357842P.
PR   21-FEB-2002; 2002US-0357843P.
PR   21-FEB-2002; 2002US-0357844P.
PR   15-MAR-2002; 2002US-0364134P.
PR   08-APR-2002; 2002US-0370144P.
PR   08-APR-2002; 2002US-0370206P.
PR   08-APR-2002; 2002US-0370247P.
PR   17-APR-2002; 2002US-0372794P.
PR   21-APR-2002; 2002US-0371679P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D,  Porter M,  Johnson K,  Higgs B,  Castle A,  Elashoff M;
```

```
XX
DR   WPI; 2003-148464/14.
XX
PT   Predicting at least one toxic effect of a compound, useful for toxicity
PT   modeling, comprises preparing a gene expression profile of a tissue or
PT   cell sample exposed to the compound, and comparing the gene expression
PT   profile to a database.
XX
PS   Example 4; Page; 446pp; English.
XX
CC   The invention relates to a novel method of predicting at least one toxic
CC   effect of a compound. The method comprises a gene expression profile of a
CC   tissue or cell sample exposed to the compound, and comparing the gene
CC   expression profile to a database comprising at least part of the data or
CC   information given in the specification. The methods are useful for
CC   predicting at least one toxic effect of a compound, predicting the
CC   progression of a toxic effect of a compound, predicting the renal
CC   toxicity of a compound, or identifying toxicity markers in tissues or
CC   cells exposed to known renal toxin. The genes are useful as toxicity
CC   markers in drug screening and toxicity assays, in monitoring disease or
CC   physiological states, or disease progression. This polynucleotide
CC   represents a rat DNA sequence relating to the toxic effect database
CC   described in the specification. NOTE: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from the World Intellectual Property
CC   Organization
XX
SQ   Sequence 4113 BP; 970 A; 1170 C; 1052 G; 921 T; 0 U; 0 Other;

  Query Match             78.2%;  Score 17.2;  DB 10;  Length 4113;
  Best Local Similarity   63.6%;  Pred. No. 1.6e+02;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||: ||: |:| ||:|||
Db       3950 GTGAACTGACTGGTGTGCTCTT 3971


RESULT 4
ADP72828
ID   ADP72828 standard; DNA; 4113 BP.
XX
AC   ADP72828;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Renal toxin progression gene marker #1417.
XX
KW   ds; toxic effect; gene expression profile; kidney tissue;
KW   differential gene expression; toxicity progression; toxicity marker;
KW   drug screening; toxicity assay; kidney pathology; nephritis;
KW   kidney necrosis; glomerular injury; tubular injury;
KW   focal segmental glomerulosclerosis.
XX
OS   Rattus norvegicus.
XX
PN   WO2004048598-A2.
XX
PD   10-JUN-2004.
XX
PF   24-NOV-2003; 2003WO-US037556.
XX
PR   22-NOV-2002; 2002US-00301856.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick DL,  Porter MW,  Johnson KR,  Castle A,  Higgs B;
PI   Elashoff M;
XX
DR   WPI; 2004-460771/43.
XX
PT   Predicting (the progression of) a toxic effect of a compound, for
PT   monitoring the progression of renal disease states, comprises preparing a
PT   gene expression profile of a kidney tissue or cell sample exposed to the
PT   compound.
XX
PS   Claim 11; SEQ ID NO 1417; 266pp; English.
XX
CC   The invention relates to a method of predicting (the progression of) a
CC   toxic effect of a compound by preparing a gene expression profile of a
CC   kidney tissue or cell sample exposed to the compound and comparing the
CC   gene expression profile to a database, or detecting the level of gene(s)
CC   expression in a tissue or cell sample exposed to the compound, where
CC   differential gene expression compared to a control indicates a toxic
CC   effect (toxicity progression). The method is useful for predicting (the
CC   progression of) at least one toxic effect of a compound. The genes are
CC   useful as toxicity markers in drug screening and toxicity assays. The
CC   methods are useful for predicting the likelihood that a compound or test
CC   agent will induce various specific kidney pathologies, such as nephritis,
CC   kidney necrosis, glomerular and tubular injury, or focal segmental
CC   glomerulosclerosis. The methods are useful for determining the similarity
CC   of a toxic response to one or more individual compounds and for
CC   predicting or elucidating the potential cellular pathways influenced,
CC   induced or modulated by the compound or test agent. The kit is useful for
CC   predicting or modelling the toxic response of a test compound, for
CC   monitoring the progression of renal disease states, for identifying genes
CC   that show promise as new drug targets and for screening known and newly
CC   designed drugs. This sequence corresponds to a gene marker used in the
CC   method of the invention. (Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences).
XX
SQ   Sequence 4113 BP; 970 A; 1170 C; 1052 G; 921 T; 0 U; 0 Other;

  Query Match             78.2%;  Score 17.2;  DB 12;  Length 4113;
  Best Local Similarity   63.6%;  Pred. No. 1.6e+02;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||: ||: |:| ||:|||
Db       3950 GTGAACTGACTGGTGTGCTCTT 3971


RESULT 5
AAV26965
ID   AAV26965 standard; cDNA to mRNA; 4131 BP.
XX
AC   AAV26965;
XX
DT   01-SEP-1998  (first entry)
XX
DE   Rat kidney calcium receptor 3A gene 4Kb fragment.
XX
KW   ss; calcium ion concentration; parathyroid hormone; homeostasis; kidney;
KW   calcium receptor; detection.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             574. .3813
FT                   /*tag= a
FT                   /product= "pRakCaR 3A 4Kb fragment"
XX
PN   US5763569-A.
XX
PD   09-JUN-1998.
XX
PF   07-JUN-1995;   95US-00484565.
XX
PR   23-AUG-1991;   91US-00749451.
PR   11-FEB-1992;   92US-00834044.
PR   21-AUG-1992;   92US-00934161.
PR   12-FEB-1993;   93US-00017127.
PR   23-FEB-1993;   93US-00009389.
PR   22-OCT-1993;   93US-00141248.
PR   19-AUG-1994;   94US-00292827.
```

```
PR   21-OCT-1994;   94WO-US012117.
PR   08-DEC-1994;   94US-00353784.
XX
PA   (NPSP-) NPS PHARM INC.
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX
PI   Hebert SC,  Brown EM,  Garrett JE;
XX
DR   WPI; 1998-347412/30.
DR   P-PSDB; AAW54847.
XX
PT   Calcium receptor poly:peptide(s) - useful for drug screening or antibody
PT   production.
XX
PS   Example 30; Fig 50; 174pp; English.
XX
CC   The rat kidney calcium receptor gene encodes a 1079 amino acid protein.
CC   The tissue from which this receptor and receptors from bovine parathyroid
CC   and rat kidney are derived, respond to changes, and control changes, in
CC   calcium ion concentration, e.g. parathyroid hormone regulates Ca2+
CC   homeostasis in blood and extracellular fluid, and kidney function alters
CC   through changes in Ca2+ levels in juxtaglomerular and proximal tubule
CC   cells in the kidney. The purified receptors (produced recombinantly) can
CC   be used to screen for compounds that modulate calcium receptor activity,
CC   especially those that can be used to treat diseases associated with the
CC   receptors in these tissues. They can also be used to raise antibodies for
CC   use in detection assays
XX
SQ   Sequence 4131 BP; 988 A; 1170 C; 1052 G; 921 T; 0 U; 0 Other;

  Query Match             78.2%;  Score 17.2;  DB 2;  Length 4131;
  Best Local Similarity   63.6%;  Pred. No. 1.6e+02;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||: ||: |:| ||:|||
Db       3950 GTGAACTGACTGGTGTGCTCTT 3971


RESULT 6
AAT95860
ID   AAT95860 standard; cDNA to mRNA; 4131 BP.
XX
AC   AAT95860;
XX
DT   08-MAY-1998  (first entry)
XX
DE   Rat kidney cell calcium receptor 3A (RakCaR 3A) cDNA.
XX
KW   Rat kidney cell calcium receptor 3A; RakCaR 3A; calcium homeostasis;
KW   hyperparathyroidism; osteoporosis; ss.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             574. .3813
FT                   /*tag=  a
FT                   /product= "kidney_cell_calcium_receptor_3A"
XX
PN   US5688938-A.
XX
PD   18-NOV-1997.
XX
PF   07-JUN-1995;   95US-00485588.
XX
PR   23-AUG-1991;   91US-00749451.
PR   11-FEB-1992;   92US-00834044.
PR   21-AUG-1992;   92US-00934161.
PR   12-FEB-1993;   93US-00017127.
PR   23-FEB-1993;   93US-00009389.
PR   22-OCT-1993;   93US-00141248.
PR   19-AUG-1994;   94US-00292827.
PR   21-OCT-1994;   94WO-US012117.
PR   08-DEC-1994;   94US-00353784.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
PA   (NPSP-) NPS PHARM INC.
XX
PI   Garrett JE,  Fuller FH,  Brown EM,  Hebert SC;
XX
DR   WPI; 1998-008040/01.
DR   P-PSDB; AAW38275.
XX
PT   DNA encoding calcium receptor polypeptide(s) - useful for therapeutic
PT   purposes, e.g. hyperparathyroidism and osteoporosis.
XX
PS   Claim 15; Col 133-142; 174pp; English.
XX
CC   The present sequence encodes rat kidney cell calcium receptor 3A (RakCaR
CC   3A). The specification includes details of molecules that can modulate
CC   one or more inorganic ion receptor activities, and antibodies and
CC   antibody fragments targetted to inorganic ion receptor proteins. The
CC   proteins, nucleic acids and antibodies may be used to treat disorders by
CC   modulating one or more inorganic ion receptor activities, preferably
CC   disorders of calcium homeostasis, e.g. hyperparathyroidism and
CC   osteoporosis
XX
SQ   Sequence 4131 BP; 987 A; 1170 C; 1053 G; 921 T; 0 U; 0 Other;

  Query Match             78.2%;  Score 17.2;  DB 2;  Length 4131;
  Best Local Similarity   63.6%;  Pred. No. 1.6e+02;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||: ||: |:| ||:|||
Db       3950 GTGAACTGACTGGTGTGCTCTT 3971


RESULT 7
AAZ25056
ID   AAZ25056 standard; cDNA to mRNA; 4131 BP.
XX
AC   AAZ25056;
XX
DT   08-DEC-1999  (first entry)
XX
DE   Rat parathyroid calcium receptor 3A nucleotide sequence.
XX
KW   Parathyroid; calcium receptor; inorganic ion receptor; modulator;
KW   receptor expression; detection; ss.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             574. .3813
FT                   /*tag=  a
FT                   /product= "RakCaR 3A"
FT                   /note= "parathyroid calcium receptor"
XX
PN   US5962314-A.
XX
PD   05-OCT-1999.
XX
PF   03-OCT-1997;   97US-00943986.
XX
PR   23-FEB-1993;   93US-00009389.
PR   22-OCT-1993;   93US-00141248.
PR   19-AUG-1994;   94US-00292827.
PR   21-OCT-1994;   94WO-US012117.
PR   08-DEC-1994;   94US-00353784.
PR   07-JUN-1995;   95US-00484565.
XX
PA   (NPSP-) NPS PHARM INC.
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX
PI   Brown EM,  Hebert SC,  Garrett JE;
```

```
XX
DR   WPI; 1999-571274/48.
DR   P-PSDB; AAY41781.
XX
PT   Nucleic acids encoding protein calcium receptors useful for identifying
PT   modulators of receptor expression and activity and for the production of
PT   antigens specific for calcium receptors.
XX
PS   Claim 8; Fig 50; 174pp; English.
XX
CC   The present sequence encodes rat parathyroid calcium receptor 3A (RakCaR
CC   3A). Calcium receptor polynucleotides may be used: (i) for producing
CC   receptor proteins (or fragments) useful for determining structure and
CC   activity relationships, for assaying molecular activity on the receptor
CC   (i.e. identifying modulators of receptor function) and for producing
CC   antibodies specific for the receptor; (ii) for sequencing the normal form
CC   of the nucleic acids (the derived sequence may be compared with other
CC   receptors to identify conserved sequences, mutations and variations that
CC   may influence calcium receptor activity and to determine target sites for
CC   antisense molecules ribozymes, hybridization probes and polymerase chain
CC   reaction (PCR) amplification primers); (iii) as hybridization probes to
CC   detect the presence of similar sequences in samples; and (iv) as PCR
CC   primers to generate particular nucleic acid sequence regions; (e.g. to
CC   generate regions to be probes by hybridization detection probes)
XX
SQ   Sequence 4131 BP; 988 A; 1170 C; 1052 G; 921 T; 0 U; 0 Other;

  Query Match             78.2%;  Score 17.2;  DB 2;  Length 4131;
  Best Local Similarity   63.6%;  Pred. No. 1.6e+02;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||: ||: |:| ||:|||
Db       3950 GTGAACTGACTGGTGTGCTCTT 3971


RESULT 8
AAV82486
ID   AAV82486 standard; cDNA to mRNA; 4131 BP.
XX
AC   AAV82486;
XX
DT   19-MAR-1999  (first entry)
XX
DE   Rat parathyroid calcium receptor pRakCaR 3A encoding cDNA.
XX
KW   Parathyroid calcium receptor; inorganic ion receptor; osteoporosis;
KW   calcium homeostasis; hyperparathyroidism; seizure; stroke; epilepsy;
KW   spinal cord injury; hypoxia-induced nerve cell damage; cardiac arrest;
KW   neonatal distress; neurodegenerative disease; Alzheimer's disease;
KW   Huntington's disease; Parkinson's disease; dementia; muscle tension;
KW   depression; anxiety; ss.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             574. .3813
FT                   /*tag=  a
XX
PN   US5858684-A.
XX
PD   12-JAN-1999.
XX
PF   07-JUN-1995;   95US-00480751.
XX
PR   23-AUG-1991;   91US-00749451.
PR   11-FEB-1992;   92US-00834044.
PR   21-AUG-1992;   92US-00934161.
PR   12-FEB-1993;   93US-00017127.
PR   23-FEB-1993;   93US-00009389.
PR   22-OCT-1993;   93US-00141248.
PR   19-AUG-1994;   94US-00292827.
PR   21-OCT-1994;   94WO-US012117.
PR   08-DEC-1994;   94US-00353784.
XX
PA   (NPSP-) NPS PHARM INC.
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX
PI   Del Mar EG,  Balandrin MF,  Van Wagenen BC,  Nemeth EF,  Brown EM;
PI   Garrett JE,  Hebert SC;
XX
DR   WPI; 1999-119871/10.
DR   P-PSDB; AAW89566.
XX
PT   Screening for calcium receptor-active compounds - by recombinant
PT   expression of nucleic acid encoding calcium receptor and determining the
PT   effect of compounds on calcium receptor activity.
XX
PS   Claim 1; Fig 50; 176pp; English.
XX
CC   A method has been developed of screening for a compound able to affect
CC   one or more activities of a calcium receptor (CR) comprises: (A)
CC   contacting a recombinant cell with a test compound, where the recombinant
CC   cell comprises a recombinant nucleic acid expressing the CR, provided
CC   that the cell does not have functional CR expression from endogenous
CC   nucleic acid; (B) determining the ability of the test compound to affect
CC   one or more activities of the calcium receptor; and (C) comparing the
CC   ability with the ability of the test compound to affect the one or more
CC   CR activities in a cell not comprising the recombinant nucleic acid. The
CC   present sequence encodes rat parathyroid CR, designated a pRakCaR 3A. The
CC   nucleic acid sequence of pRakCaR 3A can be used as part of the
CC   recombinant nucleic acid in the method described above. The compounds
CC   identified can be used to treat diseases or disorders characterised by
CC   abnormal calcium homeostasis, e.g. hyperparathyroidism, osteoporosis and
CC   other bone and mineral-related disorders. They can also be used for the
CC   treatment of diseases and disorders associated with disrupted Ca2+
CC   responses, e.g. seizures, stroke, spinal cord injury, hypoxia-induced
CC   nerve cell damage such as in cardiac arrest or neonatal distress,
CC   epilepsy, neurodegenerative diseases such as Alzheimer's disease,
CC   Huntington's disease and Parkinson's disease, dementia, muscle tension,
CC   depression, and anxiety
XX
SQ   Sequence 4131 BP; 988 A; 1170 C; 1052 G; 921 T; 0 U; 0 Other;

  Query Match             78.2%;  Score 17.2;  DB 2;  Length 4131;
  Best Local Similarity   63.6%;  Pred. No. 1.6e+02;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||: ||: |:| ||:|||
Db       3950 GTGAACTGACTGGTGTGCTCTT 3971


RESULT 9
AAZ89299
ID   AAZ89299 standard; cDNA to mRNA; 4131 BP.
XX
AC   AAZ89299;
XX
DT   09-JUN-2000  (first entry)
XX
DE   Rat calcium receptor pRakCaR3A cDNA.
XX
KW   Calcium receptor; treatment; calcimimetic; calcilytic; osteopathic;
KW   cerebroprotective; cytostatic; neuroprotective; dermatological;
KW   tranquilizer; vulnerary; antiulcer; immunosuppressive; hypotensive;
KW   cardiant; parathyroid hormone; osteoporosis; calcitonin secretion;
KW   hyperparathyroidism; Paget's disease; rat; ss.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             574. .3813
FT                   /*tag=  a
FT                   /product= "pRakCaR3A"
XX
```

PN US6031003-A.
XX
PD 29-FEB-2000.
XX
PF 07-JUN-1995; 95US-00484719.
XX
PR 23-AUG-1991; 91US-00749451.
PR 11-FEB-1992; 92US-00834044.
PR 21-AUG-1992; 92US-00934161.
PR 12-FEB-1993; 93US-00017127.
PR 23-FEB-1993; 93US-00009389.
PR 22-OCT-1993; 93US-00141248.
PR 19-AUG-1994; 94US-00292827.
PR 21-OCT-1994; 94WO-US012117.
PR 08-DEC-1994; 94US-00353784.
XX
PA (NPSP-) NPS PHARM INC.
PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
XX
PI Balandrin MF, Delmar EG, Moe ST, Nemeth EF, Van Wagenen BC;
XX
DR WPI; 2000-301969/26.
DR P-PSDB; AAY51828.
XX
PT Treating disorders or diseases in a patient by modulating inorganic ion
PT receptor activities especially calcium receptor for hyperparathyroidism,
PT by administering a calcimimetic or calcilytic compound.
XX
PS Example 30; Col 133-140; 194pp; English.
XX
CC This invention describes a novel method of treating disorders by
CC modulating calcium receptor activity in vitro by administering a
CC calcimimetic, or a calcilytic compound. The products of the invention
CC have osteopathic, cerebroprotective, cytostatic, neuroprotective,
CC dermatological, tranquilizer, vulnerary, antiulcer, immunosuppressive,
CC hypotensive and cardiant activity. The method is suitable for reducing
CC parathyroid hormone level in a patient to that of a normal individual,
CC treating a patient having osteoporosis, to inhibit bone resorption, and
CC to stimulate calcitonin secretion in vitro or in vivo. The level of
CC parathyroid hormone is reduced to cause a decrease in plasma Ca2+. The
CC method is useful in treating disorders in humans such as
CC hyperparathyroidism, Paget's disease and osteoporosis. Also for treatment
CC or prevention, based on the affected cells, of other disorders and
CC conditions like seizures, stroke, head trauma, spinal cord injury,
CC hypoxia-induced nerve cell damage such as in cardiac arrest or neonatal
CC distress, epilepsy, Alzheimer's disease, Huntington's disease,
CC Parkinson's disease, dementia, muscle tension, depression, anxiety, panic
CC disorder, OCD (not defined), post-traumatic stress disorder,
CC schizophrenia, neuroleptic malignant syndrome and Tourette's syndrome,
CC diseases involving excess water reabsorption by the kidney such as
CC syndrome of inappropriate ADH secretion (SIADH), cirrhosis, congestive
CC heart failure and nephrosis, hypertension, preventing and/or decreasing
CC renal toxicity from cationic antibiotics (e.g. aminoglycoside
CC antibiotics), gut motility disorders such as diarrhoea and spastic colon,
CC GI (gastrointestinal) ulcer diseases, GI diseases with excessive calcium
CC absorption such as sarcoidosis and autoimmune diseases and organ
CC transplant rejection. This sequence encodes the rat calcium receptor
CC pRakCaR3A which is described in the method of the invention
XX
SQ Sequence 4131 BP; 988 A; 1170 C; 1052 G; 921 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 3; Length 4131;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCGUGAGCUCTT 22
|:||||: ||: |:| ||:|||
Db 3950 GTGAACTGACTGGTGTGCTCTT 3971

RESULT 10
AAI72123
ID AAI72123 standard; cDNA to mRNA; 4131 BP.
XX
AC AAI72123;
XX
DT 25-MAR-2002 (first entry)
XX
DE cDNA encoding pRakCaR3A.
XX
KW Calcium receptor; bovine; human; parathyroid; calcium receptor; thyroid;
KW C-cell; inorganic ion receptor; homeostasis; sarcoidosis;
KW hyperparathyroidism; osteoporosis; central nervous system; seizure;
KW stroke; head trauma; spinal cord injury; organ transplant rejection;
KW hypoxia-induced nerve cell damage; cardiac arrest; neonatal distress;
KW epilepsy; neurodegenerative disease; Alzheimer's disease; cirrhosis;
KW Huntington's disease; Parkinson's disease; dementia; depression; anxiety;
KW panic disorder; obsessive-compulsive disorder; spastic colon;
KW post-traumatic stress disorder; schizophrenia; diarrhoea; kidney;
KW neuroleptic malignant syndrome; Tourette's syndrome; gut motility;
KW inappropriate ADH secretion; SIADH; gastrointestinal ulcer disease;
KW congestive heart failure; nephrosis; hypertension; alternative splicing;
KW aminoglycoside antibiotic; ss.
XX
OS Rattus rattus.
XX
FH Key Location/Qualifiers
FT CDS 574. .3813
FT /*tag= a
FT /product= "RakCaR3A"
XX
PN US6313146-B1.
XX
PD 06-NOV-2001.
XX
PF 07-JUN-1995; 95US-00484159.
XX
PR 23-AUG-1991; 91US-00749451.
PR 11-FEB-1992; 92US-00834044.
PR 21-AUG-1992; 92US-00934161.
PR 12-FEB-1993; 93US-00017127.
PR 23-FEB-1993; 93US-00009389.
PR 22-OCT-1993; 93US-00141248.
PR 19-AUG-1994; 94US-00292827.
PR 21-OCT-1994; 94WO-US012117.
PR 08-DEC-1994; 94US-00353784.
XX
PA (NPSP-) NPS PHARM INC.
XX
PI Van Wagenen BC, Balandrin MF, Delmar EG, Nemeth EF;
XX
DR WPI; 2002-081872/11.
DR P-PSDB; AAB47823.
XX
PT Novel inorganic ion receptor-modulating compounds, useful for treating
PT e.g. hyperparathyroidism, osteoporosis, stroke, epilepsy, Alzheimer's
PT disease, dementia, depression, anxiety, hypertension, cirrhosis and
PT spastic colon.
XX
PS Example 30; Fig 50; 227pp; English.
XX
CC The sequences given in AAI72120-23 encode various calcium receptor
CC proteins. This sequence encodes a rat calcium receptor. The calcium
CC receptor proteins are used, in conjunction with the compounds of the
CC invention, for structure determination, to assay a molecules activity on
CC a receptor and to obtain antibodies to that receptor. The compounds of
CC the invention, which modulate inorganic ion receptors are useful for
CC treating and diagnosing diseases or disorders which can be treated by
CC modulating inorganic ion receptor activity. This is preferably a disease
CC or disorder characterized by abnormal inorganic ion homeostasis,
CC preferably abnormal calcium homeostasis (hyperparathyroidism,
CC osteoporosis and other bone and mineral-related disorders), an abnormal
CC level of a messenger whose production or secretion is affected by
CC inorganic ion receptor activity, and an abnormal level or activity of a
CC messenger whose function is affected by inorganic ion receptor activity.
CC These compounds are also useful for treating and diagnosing diseases or

```
CC   disorders of the central nervous system such as seizures, stroke, head
CC   trauma, spinal cord injury, hypoxia-induced nerve cell damage such as in
CC   cardiac arrest or neonatal distress, epilepsy, neurodegenerative diseases
CC   such as Alzheimer's disease, Huntington's disease and Parkinson's
CC   disease, dementia, depression, anxiety, panic disorder, obsessive-
CC   compulsive disorder, post-traumatic stress disorder, schizophrenia,
CC   neuroleptic malignant syndrome and Tourette's syndrome, diseases
CC   involving excess water reabsorption by the kidney such as inappropriate
CC   ADH secretion (SIADH), cirrhosis, congestive heart failure, nephrosis,
CC   hypertension, for preventing and/or decreasing renal toxicity from
CC   cationic antibiotics (e.g. aminoglycoside antibiotics), gut motility
CC   disorders such as diarrhoea, and spastic colon, gastrointestinal (GI)
CC   ulcer diseases, GI diseases with excessive calcium absorption such as
CC   sarcoidosis, and autoimmune diseases and organ transplant rejection
XX
SQ   Sequence 4131 BP; 988 A; 1170 C; 1052 G; 921 T; 0 U; 0 Other;

  Query Match             78.2%;  Score 17.2;  DB 6;  Length 4131;
  Best Local Similarity   63.6%;  Pred. No. 1.6e+02;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||: ||: |:| ||:|||
Db       3950 GTGAACTGACTGGTGTGCTCTT 3971


RESULT 11
ADO30115
ID   ADO30115 standard; cDNA; 4550 BP.
XX
AC   ADO30115;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Mouse GPCR CASR polynucleotide, SEQ ID NO:1217.
XX
KW   G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW   transgenic mouse; neurological disorder; adrenal gland disorder;
KW   colon disorder; intestinal disorder; cardiovascular disorder;
KW   muscular disorder; blood disorder; immune disorder; bone disorder;
KW   joint disorder; metabolic disorder; nutritive disorder; cancer;
KW   kidney disorder; liver disorder; lung disorder; breast disorder;
KW   ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW   skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW   thymus disorder; thyroid disorder; antiparkinsonian; antimanic;
KW   cytostatic; antiinflammatory; vasotropic; antianginal; antiarrhythmic;
KW   CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
KW   virucide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
KW   dermatological; antiulcer; antithyroid; antiallergic; anorectic;
KW   immunosuppressive; nephrotropic; gene therapy; GPCR modulator; mouse;
KW   murine; gene; ss.
XX
OS   Mus musculus.
XX
PN   WO2004040000-A2.
XX
PD   13-MAY-2004.
XX
PF   09-SEP-2003; 2003WO-US028226.
XX
PR   09-SEP-2002; 2002US-0409303P.
PR   09-APR-2003; 2003US-0461329P.
XX
PA   (PRIM-) PRIMAL INC.
XX
PI   Gaitanaris GA,  Bergmann JE,  Gragerov A,  Hohmann J,  Li F;
PI   Madisen L,  Mcilwain KL,  Pavlova MN,  Vassilatis D,  Zeng H;
XX
DR   WPI; 2004-390329/36.
DR   P-PSDB; ADO29212.
XX
PT   Novel mammalian G protein coupled receptors, useful for identifying
PT   compounds that modulates diagnosing and treating disease condition
PT   associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT   pectoris, Parkinson's disease.
XX
PS   Claim 151; SEQ ID NO 1217; 542pp; English.
XX
CC   The invention relates to human and mouse G protein-coupled receptors
CC   (GPCRs) and nucleic acids encoding them. The invention also relates to
CC   sequences at least 90% identical to the GPCR proteins and nucleic acids
CC   of the invention; methods of treating, preventing or diagnosing diseases
CC   associated with GPCRs of the invention; methods of screening for
CC   compounds useful in the treatment of GPCR-related diseases; a transgenic
CC   mouse comprising a GPCR gene of the invention; a mouse comprising a
CC   mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC   from the trasngenic mice; kits comprising several mice, each of which has
CC   a mutation in a different GPCR gene of the invention; and kits comprising
CC   probes which hybridise to GPCR polynucleotides of the invention. The
CC   invention further discloses variants of the GPCR polypeptides and vectors
CC   comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC   be used in the diagnosis, treatment or prevention of a wide variety of
CC   diseases including neurological disorders (e.g., Alzheimer's disease,
CC   depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC   disorders of the adrenal gland; disorders of the colon or intestine
CC   (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC   syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC   myocardial infarction); muscular disorders; blood disorders (e.g.,
CC   anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC   AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC   arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC   obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC   diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC   uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC   thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
CC   nucleic acid of the invention. Note: The full sequence data for this
CC   patent did not form part of the printed specification; those sequences
CC   not shown were obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 4550 BP; 1107 A; 1287 C; 1126 G; 1030 T; 0 U; 0 Other;

  Query Match             78.2%;  Score 17.2;  DB 12;  Length 4550;
  Best Local Similarity   63.6%;  Pred. No. 1.6e+02;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||: ||: |:| ||:|||
Db       3891 GTGAACTGACTGGTGGGCTCTT 3912


RESULT 12
ACN44600
ID   ACN44600 standard; DNA; 57561 BP.
XX
AC   ACN44600;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Mouse genomic sequence mCG22161.
XX
KW   Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX
OS   Mus musculus.
XX
PN   WO2003073826-A2.
XX
PD   12-SEP-2003.
XX
PF   28-FEB-2003; 2003WO-US006235.
XX
PR   01-MAR-2002; 2002US-00087192.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW;
```

```
XX
DR   WPI; 2003-328604/31.
XX
PT   Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT   comprises a nucleotide sequence.
XX
PS   Claim 1; SEQ ID NO 1129; 0pp; English.
XX
CC   The present invention relates to novel DNA and protein sequences which
CC   are associated with carcinomas. The sequences are useful for: (i) for
CC   screening drug candidates; (ii) for screening of bioactive agent capable
CC   of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC   a bioactive agent capable of modulating the activity of CAP; (iv) for
CC   evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC   carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC   carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC   (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC   determining Carcinoma Associated (CA) gene copy number. In addition, the
CC   CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC   carcinoma including lymphoma. The present sequence is one such CA coding
CC   sequence. Note: This patent is an equivalent to basic patent
CC   US2002182586A1, for which no sequence data was published
XX
SQ   Sequence 57561 BP; 14226 A; 12066 C; 13347 G; 16427 T; 0 U; 1495 Other;

  Query Match             78.2%;  Score 17.2;  DB 11;  Length 57561;
  Best Local Similarity   63.6%;  Pred. No. 2.3e+02;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||:  |: |.:||||:|||
Db      54593 GTGAACTACCTGGTGAGCTCTT 54614


RESULT 13
ABD32923_5/c
Continuation (6 of 8) of ABD32923 from base 500001 (Human cancer-associated genomic DNA
WP Sequence split into 8 fragments  LOCUS ABD32923 Accession Abd32923
WP     Fragment Name      Begin     End
WP     ABD32923_0             1     110000
WP     ABD32923_1        100001     210000
WP     ABD32923_2        200001     310000
WP     ABD32923_3        300001     410000
WP     ABD32923_4        400001     510000
WP     ABD32923_5        500001     610000
WP     ABD32923_6        600001     710000
WP     ABD32923_7        700001     788759

  Query Match             78.2%;  Score 17.2;  DB 13;  Length 110000;
  Best Local Similarity   72.7%;  Pred. No. 2.6e+02;
  Matches   16;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              | ||||:||  ||:||||:|||
Db      15732 GAGAACTCAGCCGTGAGCTCTT 15711


RESULT 14
AAQ21468
ID   AAQ21468 standard; cDNA; 44 BP.
XX
AC   AAQ21468;
XX
DT   25-MAR-2003  (revised)
DT   02-JUN-1992  (first entry)
XX
DE   ZC2938- PCR primer for gamma-carboxylase proteins.
XX
KW   Vitamin K dependent proteins; amplification; degenerate;
KW   oligonucleotides; ss.
XX
OS   Bos taurus.
XX
PN   WO9201795-A.
XX
PD   06-FEB-1992.
XX
PF   23-JUL-1990;   90US-00557220.
XX
PR   23-JUL-1990;   90US-00557220.
PR   14-MAR-1991;   91US-00669735.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Berkner KL;
XX
DR   WPI; 1992-064951/08.
XX
PT   Gamma-carboxylase protein compsns. - used in recombinant prodn. of active
PT   vitamin=K dependent proteins.
XX
PS   Disclosure; Page 55; 91pp; English.
XX
CC   The first strand primer ZC2938 was prepared from bovine cDNA and used as
CC   a template cDNA for PCR reactions. ZC2938 (SEQ ID No 9) was synthesised
CC   from one time poly d(T)-selected bovine liver poly (A)+ RNA in two
CC   separate reactions, to assess the quality of both first and second strand
CC   synthesis. PCR was carried out by standard methods using a set of
CC   degenerate oligonucleotides complementary to the N-terminus of a gamma-
CC   carboxylase protein. Amplified DNA was ligated into a plasmid and the
CC   plasmids were selected for a gamma-carboxylase insertion. See also
CC   AAQ21467-77 (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 44 BP; 10 A; 7 C; 7 G; 20 T; 0 U; 0 Other;

  Query Match             76.4%;  Score 16.8;  DB 2;  Length 44;
  Best Local Similarity   70.0%;  Pred. No. 1.2e+02;
  Matches   14;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GAACUCACUCGUGAGCUCTT 22
              ||| :|||: |:||||:|||
Db         12 GAATTCACTAGTGAGCTCTT 31


RESULT 15
AAQ59009
ID   AAQ59009 standard; DNA; 44 BP.
XX
AC   AAQ59009;
XX
DT   25-MAR-2003  (revised)
DT   16-NOV-1994  (first entry)
XX
DE   Primer to amplify cDNA of T-47D beast carcinoma cells.
XX
KW   human calcitonin receptor; hCR; bone resorption disorders; detection;
KW   diagnosis; ss.
XX
OS   Synthetic.
XX
PN   WO9408006-A1.
XX
PD   14-APR-1994.
XX
PF   17-SEP-1993;   93WO-US008807.
XX
PR   30-SEP-1992;   92US-00954804.
PR   02-AUG-1993;   93US-00100887.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Moore EE,  Sheppard PO,  Kuestner RE;
XX
DR   WPI; 1994-135573/16.
XX
PT   DNA encoding human calcitonin receptor - used to develop drugs for
```

PT treatment and prevention of bone resorption disorders and for detection
PT and diagnosis.
XX
PS Example 1; Page 61; 83pp; English.
XX
CC T-47D human breast carcinoma cells were used to prepare a cDNA library
CC used for the cloning of calcitonin receptor sequences, AAQ59009 was used
CC in first strand cDNA synthesis from the poly(A)+ RNA. Human calcitonin
CC receptors can be used in the development of drugs treatment of bone
CC resorption disorders. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 44 BP; 10 A; 7 C; 7 G; 20 T; 0 U; 0 Other;

Query Match 76.4%; Score 16.8; DB 2; Length 44;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```
Qy          3 GAACUCACUCGUGAGCUCTT 22
              ||| :|||: |:||||:|||
Db         12 GAATTCACTAGTGAGCTCTT 31
```

Search completed: June 24, 2005, 00:22:06
Job time : 199.283 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 23, 2005, 23:58:22 ; Search time 56.4724 Seconds
(without alignments)
637.445 Million cell updates/sec

Title: US-10-848-737-1
Perfect score: 22
Sequence: 1 gugaacucacucgugagcuctt 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	17.2	78.2	4131	1	US-08-485-588-4	Sequence 4, Appli
	2	17.2	78.2	4131	1	US-08-484-565-4	Sequence 4, Appli
	3	17.2	78.2	4131	2	US-08-480-751-4	Sequence 4, Appli
	4	17.2	78.2	4131	2	US-08-943-986-4	Sequence 4, Appli
	5	17.2	78.2	4131	3	US-08-353-784-4	Sequence 4, Appli
	6	17.2	78.2	4131	3	US-08-484-719B-4	Sequence 4, Appli
	7	17.2	78.2	4131	3	US-08-484-159-4	Sequence 4, Appli
	8	16.8	76.4	44	1	US-08-361-920-38	Sequence 38, Appl
	9	16.8	76.4	44	1	US-08-453-742-15	Sequence 15, Appl
	10	16.8	76.4	44	1	US-08-454-464-15	Sequence 15, Appl
	11	16.8	76.4	44	1	US-08-453-222-15	Sequence 15, Appl
	12	16.8	76.4	44	1	US-08-452-802-15	Sequence 15, Appl
	13	16.8	76.4	44	1	US-08-479-939-38	Sequence 38, Appl
	14	16.8	76.4	44	1	US-08-483-432-38	Sequence 38, Appl
	15	16.8	76.4	44	3	US-09-071-224-12	Sequence 12, Appl
c	16	16.8	76.4	1902	4	US-09-902-540-2344	Sequence 2344, Ap
c	17	16.8	76.4	12322	4	US-09-949-016-16446	Sequence 16446, A
	18	16.8	76.4	12898	4	US-09-902-540-1000	Sequence 1000, Ap
c	19	16.8	76.4	64610	4	US-09-949-016-12214	Sequence 12214, A
	20	16.2	73.6	525	4	US-09-583-110-1469	Sequence 1469, Ap
	21	16.2	73.6	552	4	US-09-107-433-2198	Sequence 2198, Ap
c	22	16.2	73.6	1159	3	US-08-858-207A-29	Sequence 29, Appl
	23	16.2	73.6	11831	3	US-08-961-527-65	Sequence 65, Appl
c	24	16.2	73.6	13361	4	US-09-949-016-12478	Sequence 12478, A
c	25	16.2	73.6	13785	4	US-09-949-016-15631	Sequence 15631, A
	26	16.2	73.6	24428	4	US-09-949-016-17262	Sequence 17262, A
c	27	16.2	73.6	24538	4	US-09-949-016-13100	Sequence 13100, A
c	28	16.2	73.6	141115	4	US-09-949-016-17490	Sequence 17490, A
c	29	16.2	73.6	450395	4	US-09-949-016-15473	Sequence 15473, A
	30	15.8	71.8	30	2	US-08-389-423-16	Sequence 16, Appl
	31	15.8	71.8	30	3	US-09-189-028-16	Sequence 16, Appl
c	32	15.8	71.8	29598	3	US-09-341-587-6	Sequence 6, Appli
	33	15.6	70.9	601	4	US-09-949-016-58266	Sequence 58266, A
c	34	15.6	70.9	601	4	US-09-949-016-88911	Sequence 88911, A
c	35	15.6	70.9	601	4	US-09-949-016-88912	Sequence 88912, A
c	36	15.6	70.9	2433	4	US-09-300-958A-24	Sequence 24, Appl
	37	15.6	70.9	11778	4	US-09-902-540-1020	Sequence 1020, Ap
c	38	15.6	70.9	11785	1	US-08-038-768A-4	Sequence 4, Appli
c	39	15.6	70.9	11785	2	US-08-416-603-3	Sequence 3, Appli
	40	15.6	70.9	47471	4	US-09-949-016-12271	Sequence 12271, A
	41	15.6	70.9	87205	4	US-09-949-016-13430	Sequence 13430, A
	42	15.6	70.9	113283	4	US-09-949-016-16976	Sequence 16976, A
	43	15.6	70.9	113283	4	US-09-949-016-16977	Sequence 16977, A
	44	15.6	70.9	119649	4	US-09-949-016-12537	Sequence 12537, A
	45	15.6	70.9	162465	4	US-09-949-016-14264	Sequence 14264, A

ALIGNMENTS

RESULT 1
US-08-485-588-4
; Sequence 4, Application US/08485588
; Patent No. 5688938
; GENERAL INFORMATION:
; APPLICANT: Edward M. Brown
; APPLICANT: Steven C. Hebert
; APPLICANT: Forrest H. Fuller
; APPLICANT: James E. Garrett, Jr.
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,588
; FILING DATE: 7 June, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below: 9
; APPLICATION NUMBER: 08/353,784
; FILING DATE: 9 December, 1994
; APPLICATION NUMBER: PCT/US/94/12117
; FILING DATE: 21 October, 1994
; APPLICATION NUMBER: U.S. 08/292,827
; FILING DATE: 23 August, 1994
; APPLICATION NUMBER: U.S. 08/141,248
; FILING DATE: 22 October, 1993
; APPLICATION NUMBER: U.S. 08/009,389
; FILING DATE: 23 February, 1993
; APPLICATION NUMBER: U.S. 08/017,127
; FILING DATE: 12 February, 1993
; APPLICATION NUMBER: U.S. 07/934,161
; FILING DATE: 21 August, 1992
; APPLICATION NUMBER: U.S. 07/834,044
; FILING DATE: 11 February, 1992
; APPLICATION NUMBER: U.S. 07/749,451

```
;       FILING DATE:  23 August, 1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Heber, Sheldon O.
;       REGISTRATION NUMBER:  38,179
;       REFERENCE/DOCKET NUMBER:  213/005
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (213) 489-1600
;       TELEFAX:  (213) 955-0440
;       TELEX:  67-3510
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4131 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA to mRNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  574..3810
;       OTHER INFORMATION:
US-08-485-588-4

  Query Match             78.2%;  Score 17.2;  DB 1;  Length 4131;
  Best Local Similarity   63.6%;  Pred. No. 47;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||: ||: |:| ||:|||
Db       3950 GTGAACTGACTGGTGTGCTCTT 3971


RESULT 2
US-08-484-565-4
; Sequence 4, Application US/08484565
; Patent No. 5763569
;  GENERAL INFORMATION:
;    APPLICANT:  Edward M. Brown
;    APPLICANT:  Steven C. Hebert
;    APPLICANT:  James E. Garrett, Jr.
;    TITLE OF INVENTION:  CALCIUM RECEPTOR-ACTIVE
;    TITLE OF INVENTION:  MOLECULES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  First Interstate World Center
;      STREET:  Suite 4700
;      STREET:  633 West Fifth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FASTSEQ
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,565
;      FILING DATE:  7 June, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:  9
;      APPLICATION NUMBER:  08/353,784
;      FILING DATE:  9 December, 1994
;      APPLICATION NUMBER:  PCT/US/94/12117
;      FILING DATE:  21 October, 1994
;      APPLICATION NUMBER:  U.S. 08/292,827
;      FILING DATE:  23 August, 1994
;      APPLICATION NUMBER:  U.S. 08/141,248
;      FILING DATE:  22 October, 1993
;      APPLICATION NUMBER:  U.S. 08/009,389
;      FILING DATE:  23 February, 1993
;      APPLICATION NUMBER:  U.S. 08/017,127
;      FILING DATE:  12 February, 1993
;      APPLICATION NUMBER:  U.S. 07/934,161
;      FILING DATE:  21 August, 1992
;      APPLICATION NUMBER:  U.S. 07/834,044
;      FILING DATE:  11 February, 1992
;      APPLICATION NUMBER:  U.S. 07/749,451
;      FILING DATE:  23 August, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heber, Sheldon O.
;      REGISTRATION NUMBER:  38,179
;      REFERENCE/DOCKET NUMBER:  213/006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4131 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  574..3810
;      OTHER INFORMATION:
US-08-484-565-4

  Query Match             78.2%;  Score 17.2;  DB 1;  Length 4131;
  Best Local Similarity   63.6%;  Pred. No. 47;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||: ||: |:| ||:|||
Db       3950 GTGAACTGACTGGTGTGCTCTT 3971


RESULT 3
US-08-480-751-4
; Sequence 4, Application US/08480751
; Patent No. 5858684
;  GENERAL INFORMATION:
;    APPLICANT:  Edward F. Nemeth
;    APPLICANT:  Edward M. Brown
;    APPLICANT:  Steven C. Hebert
;    APPLICANT:  Forrest H. Fuller
;    APPLICANT:  James E. Garrett, Jr.
;    TITLE OF INVENTION:  CALCIUM RECEPTOR-ACTIVE
;    TITLE OF INVENTION:  MOLECULES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  First Interstate World Center
;      STREET:  Suite 4700
;      STREET:  633 West Fifth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FASTSEQ
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,751
;      FILING DATE:  7 June, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:  9
;      APPLICATION NUMBER:  08/353,784
```

```
;       FILING DATE:  9 December, 1994
;       APPLICATION NUMBER:  PCT/US/94/12117
;       FILING DATE:  21 October, 1994
;       APPLICATION NUMBER:  U.S. 08/292,827
;       FILING DATE:  23 August, 1994
;       APPLICATION NUMBER:  U.S. 08/141,248
;       FILING DATE:  22 October, 1993
;       APPLICATION NUMBER:  U.S. 08/009,389
;       FILING DATE:  23 February, 1993
;       APPLICATION NUMBER:  U.S. 08/017,127
;       FILING DATE:  12 February, 1993
;       APPLICATION NUMBER:  U.S. 07/934,161
;       FILING DATE:  21 August, 1992
;       APPLICATION NUMBER:  U.S. 07/834,044
;       FILING DATE:  11 February, 1992
;       APPLICATION NUMBER:  U.S. 07/749,451
;       FILING DATE:  23 August, 1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Heber, Sheldon O.
;       REGISTRATION NUMBER:  38,179
;       REFERENCE/DOCKET NUMBER:  213/004
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (213) 489-1600
;       TELEFAX:  (213) 955-0440
;       TELEX:  67-3510
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4131 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA to mRNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  574..3810
;       OTHER INFORMATION:
US-08-480-751-4

  Query Match             78.2%;  Score 17.2;  DB 2;  Length 4131;
  Best Local Similarity   63.6%;  Pred. No. 47;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||: ||: |:| ||:|||
Db       3950 GTGAACTGACTGGTGTGCTCTT 3971


RESULT 4
US-08-943-986-4
; Sequence 4, Application US/08943986
; Patent No. 5962314
;  GENERAL INFORMATION:
;    APPLICANT:  Edward M. Brown
;    APPLICANT:  Steven C. Hebert
;    APPLICANT:  James E. Garrett, Jr.
;    TITLE OF INVENTION:  CALCIUM RECEPTOR-ACTIVE
;    TITLE OF INVENTION:  MOLECULES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  First Interstate World Center
;      STREET:  Suite 4700
;      STREET:  633 West Fifth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FASTSEQ
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/943,986
;      FILING DATE:  03-OCT-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/484,565
;      FILING DATE:  7-June-1995
;      APPLICATION NUMBER:  08/353,784
;      FILING DATE:  9 December, 1994
;      APPLICATION NUMBER:  PCT/US/94/12117
;      FILING DATE:  21 October, 1994
;      APPLICATION NUMBER:  U.S. 08/292,827
;      FILING DATE:  23 August, 1994
;      APPLICATION NUMBER:  U.S. 08/141,248
;      FILING DATE:  22 October, 1993
;      APPLICATION NUMBER:  U.S. 08/009,389
;      FILING DATE:  23 February, 1993
;      APPLICATION NUMBER:  U.S. 08/017,127
;      FILING DATE:  12 February, 1993
;      APPLICATION NUMBER:  U.S. 07/934,161
;      FILING DATE:  21 August, 1992
;      APPLICATION NUMBER:  U.S. 07/834,044
;      FILING DATE:  11 February, 1992
;      APPLICATION NUMBER:  U.S. 07/749,451
;      FILING DATE:  23 August, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heber, Sheldon O.
;      REGISTRATION NUMBER:  38,179
;      REFERENCE/DOCKET NUMBER:  213/006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4131 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  574..3810
;      OTHER INFORMATION:
US-08-943-986-4

  Query Match             78.2%;  Score 17.2;  DB 2;  Length 4131;
  Best Local Similarity   63.6%;  Pred. No. 47;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||: ||: |:| ||:|||
Db       3950 GTGAACTGACTGGTGTGCTCTT 3971


RESULT 5
US-08-353-784-4
; Sequence 4, Application US/08353784
; Patent No. 6011068
;  GENERAL INFORMATION:
;    APPLICANT:  Edward F. Nemeth, Edward M.
;    APPLICANT:  Brown, Steven C. Hebert,
;    APPLICANT:  Bradford C. Van Wagenen, Manuel
;    APPLICANT:  F. Balandrin, Forrest H. Fuller,
;    APPLICANT:  Eric G. DelMar, and Scott T. Moe
;    TITLE OF INVENTION:  CALCIUM RECEPTOR-ACTIVE
;    TITLE OF INVENTION:  MOLECULES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  First Interstate World Center
;      STREET:  Suite 4700
;      STREET:  633 West Fifth Street
;      CITY:  Los Angeles
```

```
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FASTSEQ
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/353,784
;      FILING DATE:  9 December, 1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:  8
;      APPLICATION NUMBER:  PCT/US/94/12117
;      FILING DATE:  21 October, 1994
;      APPLICATION NUMBER:  U.S. 08/292,827
;      FILING DATE:  23 August, 1994
;      APPLICATION NUMBER:  U.S. 08/141,248
;      FILING DATE:  22 October, 1993
;      APPLICATION NUMBER:  U.S. 08/009,389
;      FILING DATE:  23 February, 1993
;      APPLICATION NUMBER:  U.S. 08/017,127
;      FILING DATE:  12 February, 1993
;      APPLICATION NUMBER:  U.S. 07/934,161
;      FILING DATE:  21 August, 1992
;      APPLICATION NUMBER:  U.S. 07/834,044
;      FILING DATE:  11 February, 1992
;      APPLICATION NUMBER:  U.S. 07/749,451
;      FILING DATE:  23 August, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heber, Sheldon O.
;      REGISTRATION NUMBER:  38,179
;      REFERENCE/DOCKET NUMBER:  209/069
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4131 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  574..3810
;      OTHER INFORMATION:
US-08-353-784-4

  Query Match             78.2%;  Score 17.2;  DB 3;  Length 4131;
  Best Local Similarity   63.6%;  Pred. No. 47;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||: ||: |:| ||:|||
Db       3950 GTGAACTGACTGGTGTGCTCTT 3971


RESULT 6
US-08-484-719B-4
; Sequence 4, Application US/08484719B
; Patent No. 6031003
;  GENERAL INFORMATION:
;    APPLICANT:  Edward F. Nemeth, Edward M.
;    APPLICANT:  Brown, Steven C. Hebert,
;    APPLICANT:  Bradford C. Van Wagenen,
;    APPLICANT:  Manuel F. Balandrin,
;    APPLICANT:  Forrest H. Fuller, Eric G.
;    APPLICANT:  Delmar, Scott T. Moe
;    TITLE OF INVENTION:  CALCIUM RECEPTOR-ACTIVE
;    TITLE OF INVENTION:  MOLECULES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  First Interstate World Center
;      STREET:  Suite 4700
;      STREET:  633 West Fifth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS Word
;      SOFTWARE:  FastSEQ for Windows Version 3.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,719B
;      FILING DATE:  7 June, 1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/353,784
;      FILING DATE:  9 December, 1994
;      APPLICATION NUMBER:  PCT/US/94/12117
;      FILING DATE:  21 October, 1994
;      APPLICATION NUMBER:  U.S. 08/292,827
;      FILING DATE:  23 August, 1994
;      APPLICATION NUMBER:  U.S. 08/141,248
;      FILING DATE:  22 October, 1993
;      APPLICATION NUMBER:  U.S. 08/009,389
;      FILING DATE:  23 February, 1993
;      APPLICATION NUMBER:  U.S. 08/017,127
;      FILING DATE:  12 February, 1993
;      APPLICATION NUMBER:  U.S. 07/934,161
;      FILING DATE:  21 August, 1992
;      APPLICATION NUMBER:  U.S. 07/834,044
;      FILING DATE:  11 February, 1992
;      APPLICATION NUMBER:  U.S. 07/749,451
;      FILING DATE:  23 August, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Douglas C. Murdock
;      REGISTRATION NUMBER:  37,549
;      REFERENCE/DOCKET NUMBER:  213/007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4131 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  574..3810
US-08-484-719B-4

  Query Match             78.2%;  Score 17.2;  DB 3;  Length 4131;
  Best Local Similarity   63.6%;  Pred. No. 47;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||: ||: |:| ||:|||
Db       3950 GTGAACTGACTGGTGTGCTCTT 3971


RESULT 7
US-08-484-159-4
; Sequence 4, Application US/08484159
; Patent No. 6313146
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Bradford C. Van Wagenen
;    APPLICANT:  Manuel F. Balandrin
;    APPLICANT:  Eric G. Del Mar
;    APPLICANT:  Edward F. Nemeth
;    TITLE OF INVENTION:  CALCIUM RECEPTOR-ACTIVE
;    TITLE OF INVENTION:  MOLECULES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  First Interstate World Center
;      STREET:  Suite 4700
;      STREET:  633 West Fifth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FASTSEQ
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,159
;      FILING DATE:  7 June, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:  9
;      APPLICATION NUMBER:  08/353,784
;      FILING DATE:  9 December, 1994
;      APPLICATION NUMBER:  PCT/US/94/12117
;      FILING DATE:  21 October, 1994
;      APPLICATION NUMBER:  U.S. 08/292,827
;      FILING DATE:  23 August, 1994
;      APPLICATION NUMBER:  U.S. 08/141,248
;      FILING DATE:  22 October, 1993
;      APPLICATION NUMBER:  U.S. 08/009,389
;      FILING DATE:  23 February, 1993
;      APPLICATION NUMBER:  U.S. 08/017,127
;      FILING DATE:  12 February, 1993
;      APPLICATION NUMBER:  U.S. 07/934,161
;      FILING DATE:  21 August, 1992
;      APPLICATION NUMBER:  U.S. 07/834,044
;      FILING DATE:  11 February, 1992
;      APPLICATION NUMBER:  U.S. 07/749,451
;      FILING DATE:  23 August, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heber, Sheldon O.
;      REGISTRATION NUMBER:  38,179
;      REFERENCE/DOCKET NUMBER:  214/101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4131 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  574..3810
;      OTHER INFORMATION:
US-08-484-159-4

  Query Match             78.2%;  Score 17.2;  DB 3;  Length 4131;
  Best Local Similarity   63.6%;  Pred. No. 47;
  Matches   14;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||: ||: |:| ||:|||
Db       3950 GTGAACTGACTGGTGTGCTCTT 3971


RESULT 8
US-08-361-920-38
; Sequence 38, Application US/08361920
; Patent No. 5457046
;  GENERAL INFORMATION:
;    APPLICANT:  Woeldike, Helle F.
;    APPLICANT:  Hagen, Frederick
;    APPLICANT:  Hjort, Carsten M.
;    APPLICANT:  Sven, Hastrup
;    TITLE OF INVENTION:  An Enzyme Capable of Degrading Cellulose
;    TITLE OF INVENTION:  or Hemicellulose
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5457046o No. 5457046disk of No. 5457046th America, Inc.
;      STREET:  405 Lexington Avenue, 62nd Floor
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  10174-6201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/361,920
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/940,860
;      FILING DATE:  28-OCT-1992
;      APPLICATION NUMBER:  DK 1158/90
;      FILING DATE:  09-MAY-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/DK91/00124
;      FILING DATE:  08-MAY-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J.
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  3435.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-867-0298
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-361-920-38

  Query Match             76.4%;  Score 16.8;  DB 1;  Length 44;
  Best Local Similarity   70.0%;  Pred. No. 43;
  Matches   14;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GAACUCACUCGUGAGCUCTT 22
              ||| :|||: |:||||:|||
Db         12 GAATTCACTAGTGAGCTCTT 31


RESULT 9
US-08-453-742-15
; Sequence 15, Application US/08453742
; Patent No. 5622839
;  GENERAL INFORMATION:
;    APPLICANT:  Moore, Emma E
;    APPLICANT:  Sheppard, Paul O
;    APPLICANT:  Kuestner, Rolf E
;    TITLE OF INVENTION:  Human Calcitonin Receptor
```

```
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND AND TOWNSEND KHOURIE and CREW
;      STREET:  One Market Plaza, Steuart St. Tower,
;      STREET:  Twentieth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/453,742
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/100,887
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/954,804
;      FILING DATE:  30-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parmelee, Steven W
;      REGISTRATION NUMBER:  31,990
;      REFERENCE/DOCKET NUMBER:  13952-15-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-467-9600
;      TELEFAX:  415-543-5043
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  ZC2938
US-08-453-742-15

  Query Match             76.4%;  Score 16.8;  DB 1;  Length 44;
  Best Local Similarity   70.0%;  Pred. No. 43;
  Matches   14;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GAACUCACUCGUGAGCUCTT 22
              ||| :|||: |:||||:|||
Db         12 GAATTCACTAGTGAGCTCTT 31


RESULT 10
US-08-454-464-15
; Sequence 15, Application US/08454464
; Patent No. 5674689
;  GENERAL INFORMATION:
;    APPLICANT:  Moore, Emma E
;    APPLICANT:  Sheppard, Paul O
;    APPLICANT:  Kuestner, Rolf E
;    TITLE OF INVENTION:  Human Calcitonin Receptor
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND AND TOWNSEND KHOURIE and CREW
;      STREET:  One Market Plaza, Steuart St. Tower,
;      STREET:  Twentieth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,464
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/100,887
;      FILING DATE:  02-AUG-1993
;      APPLICATION NUMBER:  US 07/954,804
;      FILING DATE:  30-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parmelee, Steven W
;      REGISTRATION NUMBER:  31,990
;      REFERENCE/DOCKET NUMBER:  13952-15-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-467-9600
;      TELEFAX:  415-543-5043
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  ZC2938
US-08-454-464-15

  Query Match             76.4%;  Score 16.8;  DB 1;  Length 44;
  Best Local Similarity   70.0%;  Pred. No. 43;
  Matches   14;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GAACUCACUCGUGAGCUCTT 22
              ||| :|||: |:||||:|||
Db         12 GAATTCACTAGTGAGCTCTT 31


RESULT 11
US-08-453-222-15
; Sequence 15, Application US/08453222
; Patent No. 5674981
;  GENERAL INFORMATION:
;    APPLICANT:  Moore, Emma E
;    APPLICANT:  Sheppard, Paul O
;    APPLICANT:  Kuestner, Rolf E
;    TITLE OF INVENTION:  Human Calcitonin Receptor
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND AND TOWNSEND KHOURIE and CREW
;      STREET:  One Market Plaza, Steuart St. Tower,
;      STREET:  Twentieth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/453,222
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/100,887
;      FILING DATE:  02-AUG-1993
;      APPLICATION NUMBER:  US 07/954,804
;      FILING DATE:  30-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parmelee, Steven W
;      REGISTRATION NUMBER:  31,990
;      REFERENCE/DOCKET NUMBER:  13952-15-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-467-9600
;      TELEFAX:  415-543-5043
```

```
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  ZC2938
US-08-453-222-15

  Query Match             76.4%;  Score 16.8;  DB 1;  Length 44;
  Best Local Similarity   70.0%;  Pred. No. 43;
  Matches   14;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GAACUCACUCGUGAGCUCTT 22
              ||| :|||: |:||||:|||
Db         12 GAATTCACTAGTGAGCTCTT 31


RESULT 12
US-08-452-802-15
; Sequence 15, Application US/08452802
; Patent No. 5683884
;  GENERAL INFORMATION:
;    APPLICANT:  Moore, Emma E
;    APPLICANT:  Sheppard, Paul O
;    APPLICANT:  Kuestner, Rolf E
;    TITLE OF INVENTION:  Human Calcitonin Receptor
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND AND TOWNSEND KHOURIE and CREW
;      STREET:  One Market Plaza, Steuart St. Tower,
;      STREET:  Twentieth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/452,802
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/100,887
;      FILING DATE:  02-AUG-1993
;      APPLICATION NUMBER:  US 07/954,804
;      FILING DATE:  30-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parmelee, Steven W
;      REGISTRATION NUMBER:  31,990
;      REFERENCE/DOCKET NUMBER:  13952-15-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-467-9600
;      TELEFAX:  415-543-5043
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  ZC2938
US-08-452-802-15

  Query Match             76.4%;  Score 16.8;  DB 1;  Length 44;
  Best Local Similarity   70.0%;  Pred. No. 43;
  Matches   14;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GAACUCACUCGUGAGCUCTT 22
              ||| :|||: |:||||:|||
Db         12 GAATTCACTAGTGAGCTCTT 31


RESULT 13
US-08-479-939-38
; Sequence 38, Application US/08479939
; Patent No. 5686593
;  GENERAL INFORMATION:
;    APPLICANT:  Woeldike, Helle F.
;    APPLICANT:  Hagen, Frederick
;    APPLICANT:  Hjort, Carsten M.
;    APPLICANT:  Sven, Hastrup
;    TITLE OF INVENTION:  An Enzyme Capable of Degrading Cellulose
;    TITLE OF INVENTION:  or Hemicellulose
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5686593o No. 5686593disk of No. 5686593th America, Inc.
;      STREET:  405 Lexington Avenue, 62nd Floor
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  10174-6201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/479,939
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/361,920
;      FILING DATE:  22-DEC-1994
;      APPLICATION NUMBER:  US 07/940,860
;      FILING DATE:  28-OCT-1992
;      APPLICATION NUMBER:  DK 1158/90
;      FILING DATE:  09-MAY-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/DK91/00124
;      FILING DATE:  08-MAY-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J.
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  3435.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-867-0298
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-479-939-38

  Query Match             76.4%;  Score 16.8;  DB 1;  Length 44;
  Best Local Similarity   70.0%;  Pred. No. 43;
  Matches   14;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GAACUCACUCGUGAGCUCTT 22
              ||| :|||: |:||||:|||
Db         12 GAATTCACTAGTGAGCTCTT 31


RESULT 14
US-08-483-432-38
; Sequence 38, Application US/08483432
; Patent No. 5763254
;  GENERAL INFORMATION:
;    APPLICANT:  Woeldike, Helle F.
```

```
;    APPLICANT:  Hagen, Frederick
;    APPLICANT:  Hjort, Carsten M.
;    APPLICANT:  Sven, Hastrup
;    TITLE OF INVENTION:  An Enzyme Capable of Degrading Cellulose
;    TITLE OF INVENTION:  or Hemicellulose
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5763254o No. 5763254disk of No. 5763254th America, Inc.
;      STREET:  405 Lexington Avenue, 62nd Floor
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  10174-6201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,432
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/361,920
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/940,860
;      FILING DATE:  28-OCT-1992
;      APPLICATION NUMBER:  DK 1158/90
;      FILING DATE:  09-MAY-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/DK91/00124
;      FILING DATE:  08-MAY-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J.
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  3435.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-867-0298
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-483-432-38

  Query Match             76.4%;  Score 16.8;  DB 1;  Length 44;
  Best Local Similarity   70.0%;  Pred. No. 43;
  Matches   14;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GAACUCACUCGUGAGCUCTT 22
              ||| :|||: |:||||:|||
Db         12 GAATTCACTAGTGAGCTCTT 31


RESULT 15
US-09-071-224-12
; Sequence 12, Application US/09071224
; Patent No. 6271343
;  GENERAL INFORMATION:
;    APPLICANT:  Lok, Si
;    APPLICANT:  Presnell, Scott R.
;    APPLICANT:  Jelmberg, Anna C.
;    APPLICANT:  Gilbert, Teresa
;    APPLICANT:  Foster, Donald C.
;    APPLICANT:  Adams, Robyn L.
;    APPLICANT:  Lehner, Joyce M.
;    TITLE OF INVENTION:  MAMMALIAN ZCYTOR5
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Zymogenetics
;      STREET:  1201 Eastlake Ave East
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/071,224
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lunn, Paul G
;      REGISTRATION NUMBER:  32,743
;      REFERENCE/DOCKET NUMBER:  96-22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-442-6627
;      TELEFAX:  206-442-6678
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other
US-09-071-224-12

  Query Match             76.4%;  Score 16.8;  DB 3;  Length 44;
  Best Local Similarity   70.0%;  Pred. No. 43;
  Matches   14;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GAACUCACUCGUGAGCUCTT 22
              ||| :|||: |:||||:|||
Db         12 GAATTCACTAGTGAGCTCTT 31


Search completed: June 24, 2005, 04:11:08
Job time : 64.4724 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 24, 2005, 00:03:32 ; Search time 253.433 Seconds
(without alignments)
542.594 Million cell updates/sec

Title: US-10-848-737-1
Perfect score: 22
Sequence: 1 gugaacucacucgugagcuctt 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6067389 seqs, 3125258755 residues

Total number of hits satisfying chosen parameters: 12134778

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	22	100.0	22	21	US-10-848-737-1	Sequence 1, Appli
c 2	20	90.9	20	21	US-10-831-901A-780	Sequence 780, App
3	20	90.9	1215	21	US-10-831-901A-29766	Sequence 29766, A
4	20	90.9	1706	21	US-10-699-936-4	Sequence 4, Appli
5	20	90.9	1706	21	US-10-699-936-14	Sequence 14, Appl
6	20	90.9	24774	21	US-10-889-447-3	Sequence 3, Appli
7	20	90.9	24774	21	US-10-831-901A-29748	Sequence 29748, A
8	20	90.9	24774	21	US-10-889-101-3	Sequence 3, Appli
9	20	90.9	28920	21	US-10-889-447-5	Sequence 5, Appli
10	20	90.9	28920	21	US-10-889-447-6	Sequence 6, Appli
11	20	90.9	28920	21	US-10-831-901A-29740	Sequence 29740, A
12	20	90.9	28920	21	US-10-889-101-5	Sequence 5, Appli
13	20	90.9	28920	21	US-10-889-101-6	Sequence 6, Appli
14	20	90.9	29013	21	US-10-831-901A-29819	Sequence 29819, A
15	20	90.9	29206	21	US-10-831-901A-29742	Sequence 29742, A
16	20	90.9	29291	21	US-10-889-447-4	Sequence 4, Appli
17	20	90.9	29291	21	US-10-831-901A-29738	Sequence 29738, A
18	20	90.9	29291	21	US-10-889-101-4	Sequence 4, Appli
19	20	90.9	29429	21	US-10-831-901A-29739	Sequence 29739, A
20	20	90.9	29430	21	US-10-889-447-7	Sequence 7, Appli
21	20	90.9	29430	21	US-10-831-901A-29741	Sequence 29741, A
22	20	90.9	29430	21	US-10-889-101-7	Sequence 7, Appli
23	20	90.9	29573	21	US-10-831-901A-29802	Sequence 29802, A
24	20	90.9	29573	21	US-10-831-901A-29803	Sequence 29803, A
25	20	90.9	29573	21	US-10-831-901A-29807	Sequence 29807, A
26	20	90.9	29592	21	US-10-831-901A-29820	Sequence 29820, A
27	20	90.9	29705	21	US-10-831-901A-29758	Sequence 29758, A
28	20	90.9	29705	21	US-10-831-901A-29791	Sequence 29791, A
29	20	90.9	29706	21	US-10-831-901A-29756	Sequence 29756, A
30	20	90.9	29706	21	US-10-831-901A-29793	Sequence 29793, A
31	20	90.9	29711	21	US-10-831-901A-29755	Sequence 29755, A
32	20	90.9	29711	21	US-10-831-901A-29757	Sequence 29757, A
33	20	90.9	29711	21	US-10-831-901A-29759	Sequence 29759, A
34	20	90.9	29711	21	US-10-831-901A-29779	Sequence 29779, A
35	20	90.9	29711	21	US-10-831-901A-29790	Sequence 29790, A
36	20	90.9	29711	21	US-10-831-901A-29792	Sequence 29792, A
37	20	90.9	29711	21	US-10-831-901A-29794	Sequence 29794, A
38	20	90.9	29711	21	US-10-831-901A-29815	Sequence 29815, A
39	20	90.9	29715	21	US-10-831-901A-29760	Sequence 29760, A
40	20	90.9	29715	21	US-10-831-901A-29795	Sequence 29795, A
41	20	90.9	29715	21	US-10-831-901A-29816	Sequence 29816, A
42	20	90.9	29720	21	US-10-831-901A-29798	Sequence 29798, A
43	20	90.9	29725	21	US-10-831-901A-29753	Sequence 29753, A
44	20	90.9	29725	21	US-10-831-901A-29774	Sequence 29774, A
45	20	90.9	29725	21	US-10-831-901A-29781	Sequence 29781, A

ALIGNMENTS

RESULT 1
US-10-848-737-1
; Sequence 1, Application US/10848737
; Publication No. US20050004063A1
; GENERAL INFORMATION:
; APPLICANT: HE, MING-LIANG
; APPLICANT: KUNG, HSIANG-FU
; APPLICANT: ZHENG, BOJIAN
; APPLICANT: LIN, MARIE C. M.
; APPLICANT: PENG, YING
; APPLICANT: GUAN, YI
; TITLE OF INVENTION: INHIBITION OF SARS-ASSOCIATED CORONAVIRUS (SCoV)
; TITLE OF INVENTION: INFECTION AND REPLICATION BY RNA INTERFERENCE
; FILE REFERENCE: V9661.0080
; CURRENT APPLICATION NUMBER: US/10/848,737
; CURRENT FILING DATE: 2004-05-19
; PRIOR APPLICATION NUMBER: 60/471,901
; PRIOR FILING DATE: 2003-05-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 1
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule:
; OTHER INFORMATION: Synthetic oligonucleotide SARSi-1
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: oligonucleotide

US-10-848-737-1

Query Match 100.0%; Score 22; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              ||||||||||||||||||||||
Db          1 GUGAACUCACUCGUGAGCUCTT 22
```

RESULT 2
US-10-831-901A-780/c
; Sequence 780, Application US/10831901A
; Publication No. US20050100885A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Ecker, David J.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Freier, Susan M.
; APPLICANT: Massire, Christian
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Lowery, Kristin Sannes
; APPLICANT: Swayze, Eric
; APPLICANT: Baker, Brenda F.
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
; TITLE OF INVENTION: Acute Respiratory Syndrome (SARS)
; FILE REFERENCE: ISIS0083-100 (BIOL0008US)
; CURRENT APPLICATION NUMBER: US/10/831,901A
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/466,426
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: 60/468,562
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/467,770
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 60/468,627
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/477,637
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: 60/483,579
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 30063
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 780
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense compound
US-10-831-901A-780

Query Match 90.9%; Score 20; DB 21; Length 20;
Best Local Similarity 75.0%; Pred. No. 3.8;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db         20 GTGAACTCACTCGTGAGCTC 1
```

RESULT 3
US-10-831-901A-29766
; Sequence 29766, Application US/10831901A
; Publication No. US20050100885A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Ecker, David J.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Freier, Susan M.
; APPLICANT: Massire, Christian
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Lowery, Kristin Sannes
; APPLICANT: Swayze, Eric
; APPLICANT: Baker, Brenda F.
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
; TITLE OF INVENTION: Acute Respiratory Syndrome (SARS)
; FILE REFERENCE: ISIS0083-100 (BIOL0008US)
; CURRENT APPLICATION NUMBER: US/10/831,901A
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/466,426
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: 60/468,562
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/467,770
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 60/468,627
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/477,637
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: 60/483,579
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 30063
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29766
; LENGTH: 1215
; TYPE: DNA
; ORGANISM: SARS Coronavirus
US-10-831-901A-29766

Query Match 90.9%; Score 20; DB 21; Length 1215;
Best Local Similarity 75.0%; Pred. No. 4.2;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        760 GTGAACTCACTCGTGAGCTC 779
```

RESULT 4
US-10-699-936-4
; Sequence 4, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David E.
; APPLICANT: Jollick, Joseph D.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; TITLE OF INVENTION: Syndrome Coronavirus
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 1706
; TYPE: DNA
; ORGANISM: SARS-CoV Shanghai LY
US-10-699-936-4

Query Match 90.9%; Score 20; DB 21; Length 1706;
Best Local Similarity 75.0%; Pred. No. 4.3;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        697 GTGAACTCACTCGTGAGCTC 716
```

RESULT 5
US-10-699-936-14
; Sequence 14, Application US/10699936

```
; Publication No. US20050095582A1
; GENERAL INFORMATION:
;  APPLICANT: Gillim-Ross, Laura
;  APPLICANT:  Taylor, Jill
;  APPLICANT:  Scholl, David R.
;  APPLICANT:  Wentworth, David E.
;  APPLICANT:  Jollick, Joseph D.
;  TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
;  TITLE OF INVENTION:  Syndrome Coronavirus
;  FILE REFERENCE: DHI-07986
;  CURRENT APPLICATION NUMBER: US/10/699,936
;  CURRENT FILING DATE:  2003-11-03
;  NUMBER OF SEQ ID NOS: 87
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
;   LENGTH: 1706
;   TYPE: DNA
;   ORGANISM: SARS coronavirus Shanghai LY
US-10-699-936-14

  Query Match             90.9%;  Score 20;  DB 21;  Length 1706;
  Best Local Similarity   75.0%;  Pred. No. 4.3;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        697 GTGAACTCACTCGTGAGCTC 716


RESULT 6
US-10-889-447-3
; Sequence 3, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
;  APPLICANT: Bennett, C. Frank
;  APPLICANT:  Jain, Ravi
;  TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
;  FILE REFERENCE: RTS-0685US
;  CURRENT APPLICATION NUMBER: US/10/889,447
;  CURRENT FILING DATE:  2004-07-12
;  PRIOR APPLICATION NUMBER: 60/486,670
;  PRIOR FILING DATE: 2003-07-12
;  NUMBER OF SEQ ID NOS: 241
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 24774
;   TYPE: DNA
;   ORGANISM: SARS coronavirus isolate BJ01
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION:
;   OTHER INFORMATION: n is any nucleotide
US-10-889-447-3

  Query Match             90.9%;  Score 20;  DB 21;  Length 24774;
  Best Local Similarity   75.0%;  Pred. No. 4.6;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        706 GTGAACTCACTCGTGAGCTC 725


RESULT 7
US-10-831-901A-29748
; Sequence 29748, Application US/10831901A
; Publication No. US20050100885A1
; GENERAL INFORMATION:
;  APPLICANT: Crooke, Stanley T.
;  APPLICANT:  Ecker, David J.
;  APPLICANT:  Sampath, Rangarajan
;  APPLICANT:  Freier, Susan M.
;  APPLICANT:  Massire, Christian
;  APPLICANT:  Hofstadler, Steven A.
;  APPLICANT:  Lowery, Kristin Sannes
;  APPLICANT:  Swayze, Eric
;  APPLICANT:  Baker, Brenda F.
;  APPLICANT:  Bennett, C. Frank
;  TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
;  TITLE OF INVENTION:  Acute Respiratory Syndrome (SARS)
;  FILE REFERENCE: ISIS0083-100 (BIOL0008US)
;  CURRENT APPLICATION NUMBER: US/10/831,901A
;  CURRENT FILING DATE:  2004-04-26
;  PRIOR APPLICATION NUMBER: 60/466,426
;  PRIOR FILING DATE: 2003-04-28
;  PRIOR APPLICATION NUMBER: 60/468,562
;  PRIOR FILING DATE: 2003-05-06
;  PRIOR APPLICATION NUMBER: 60/467,770
;  PRIOR FILING DATE: 2003-04-30
;  PRIOR APPLICATION NUMBER: 60/468,627
;  PRIOR FILING DATE: 2003-05-06
;  PRIOR APPLICATION NUMBER: 60/477,637
;  PRIOR FILING DATE: 2003-06-10
;  PRIOR APPLICATION NUMBER: 60/483,579
;  PRIOR FILING DATE: 2003-06-27
;  NUMBER OF SEQ ID NOS: 30063
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29748
;   LENGTH: 24774
;   TYPE: DNA
;   ORGANISM: SARS Coronavirus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 860-959, 2089-2188, 2478-2577, 2877-2976, 3576-3675, 3865-3964,
;   LOCATION: 4134-4233, 4563-4662, 5012-5111, 5781-5879, 7438-7537, 7837-7936,
;   LOCATION: 8616-8715, 12025-12124, 13984-14083, 16463-16562, 16932-17031,
;   LOCATION: 17381-17480, 18090-18189, 19019-19118, 19478-19577, 20357-20456,
;   LOCATION: 21086-21185, 21945-22044, 23174-23273, 23531
;   OTHER INFORMATION: n = A,T,C or G
US-10-831-901A-29748

  Query Match             90.9%;  Score 20;  DB 21;  Length 24774;
  Best Local Similarity   75.0%;  Pred. No. 4.6;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        706 GTGAACTCACTCGTGAGCTC 725


RESULT 8
US-10-889-101-3
; Sequence 3, Application US/10889101
; Publication No. US20050107324A1
; GENERAL INFORMATION:
;  APPLICANT: Bennett, C. Frank
;  APPLICANT:  Dobie, Kenneth W.
;  APPLICANT:  Jain, Ravi
;  TITLE OF INVENTION: MODULATION OF CEACAM1 EXPRESSION
;  FILE REFERENCE: ISIS0101-100 (RTS-0655US)
;  CURRENT APPLICATION NUMBER: US/10/889,101
;  CURRENT FILING DATE:  2004-07-12
;  PRIOR APPLICATION NUMBER: 60/486,652
;  PRIOR FILING DATE: 2003-07-12
;  NUMBER OF SEQ ID NOS: 298
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 24774
;   TYPE: DNA
;   ORGANISM: SARS Coronavirus isolate BJ01
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION:
;   OTHER INFORMATION: n = A,T,C or G
US-10-889-101-3
```

```
  Query Match             90.9%;  Score 20;  DB 21;  Length 24774;
  Best Local Similarity   75.0%;  Pred. No. 4.6;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        706 GTGAACTCACTCGTGAGCTC 725


RESULT 9
US-10-889-447-5
; Sequence 5, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
;  APPLICANT: Bennett, C. Frank
;  APPLICANT:  Jain, Ravi
;  TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
;  FILE REFERENCE: RTS-0685US
;  CURRENT APPLICATION NUMBER: US/10/889,447
;  CURRENT FILING DATE:  2004-07-12
;  PRIOR APPLICATION NUMBER: 60/486,670
;  PRIOR FILING DATE: 2003-07-12
;  NUMBER OF SEQ ID NOS: 241
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 28920
;   TYPE: DNA
;   ORGANISM: SARS coronavirus isolate BJ03
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION:
;   OTHER INFORMATION: n is any nucleotide
US-10-889-447-5

  Query Match             90.9%;  Score 20;  DB 21;  Length 28920;
  Best Local Similarity   75.0%;  Pred. No. 4.6;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        776 GTGAACTCACTCGTGAGCTC 795


RESULT 10
US-10-889-447-6
; Sequence 6, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
;  APPLICANT: Bennett, C. Frank
;  APPLICANT:  Jain, Ravi
;  TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
;  FILE REFERENCE: RTS-0685US
;  CURRENT APPLICATION NUMBER: US/10/889,447
;  CURRENT FILING DATE:  2004-07-12
;  PRIOR APPLICATION NUMBER: 60/486,670
;  PRIOR FILING DATE: 2003-07-12
;  NUMBER OF SEQ ID NOS: 241
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 28920
;   TYPE: DNA
;   ORGANISM: SARS coronavirus isolate BJ04
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION:
;   OTHER INFORMATION: n is any nucleotide
US-10-889-447-6

  Query Match             90.9%;  Score 20;  DB 21;  Length 28920;
  Best Local Similarity   75.0%;  Pred. No. 4.6;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        776 GTGAACTCACTCGTGAGCTC 795


RESULT 11
US-10-831-901A-29740
; Sequence 29740, Application US/10831901A
; Publication No. US20050100885A1
; GENERAL INFORMATION:
;  APPLICANT: Crooke, Stanley T.
;  APPLICANT:  Ecker, David J.
;  APPLICANT:  Sampath, Rangarajan
;  APPLICANT:  Freier, Susan M.
;  APPLICANT:  Massire, Christian
;  APPLICANT:  Hofstadler, Steven A.
;  APPLICANT:  Lowery, Kristin Sannes
;  APPLICANT:  Swayze, Eric
;  APPLICANT:  Baker, Brenda F.
;  APPLICANT:  Bennett, C. Frank
;  TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
;  TITLE OF INVENTION:  Acute Respiratory Syndrome (SARS)
;  FILE REFERENCE: ISIS0083-100 (BIOL0008US)
;  CURRENT APPLICATION NUMBER: US/10/831,901A
;  CURRENT FILING DATE:  2004-04-26
;  PRIOR APPLICATION NUMBER: 60/466,426
;  PRIOR FILING DATE: 2003-04-28
;  PRIOR APPLICATION NUMBER: 60/468,562
;  PRIOR FILING DATE: 2003-05-06
;  PRIOR APPLICATION NUMBER: 60/467,770
;  PRIOR FILING DATE: 2003-04-30
;  PRIOR APPLICATION NUMBER: 60/468,627
;  PRIOR FILING DATE: 2003-05-06
;  PRIOR APPLICATION NUMBER: 60/477,637
;  PRIOR FILING DATE: 2003-06-10
;  PRIOR APPLICATION NUMBER: 60/483,579
;  PRIOR FILING DATE: 2003-06-27
;  NUMBER OF SEQ ID NOS: 30063
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29740
;   LENGTH: 28920
;   TYPE: DNA
;   ORGANISM: SARS Coronavirus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 7230-7329, 9929-10028, 10137-10236, 11646-11745, 20385-20484,
;   LOCATION: 21024-21123, 21753-21852, 22112-22211, 25301-25400
;   OTHER INFORMATION: n = A,T,C or G
US-10-831-901A-29740

  Query Match             90.9%;  Score 20;  DB 21;  Length 28920;
  Best Local Similarity   75.0%;  Pred. No. 4.6;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        776 GTGAACTCACTCGTGAGCTC 795


RESULT 12
US-10-889-101-5
; Sequence 5, Application US/10889101
; Publication No. US20050107324A1
; GENERAL INFORMATION:
;  APPLICANT: Bennett, C. Frank
;  APPLICANT:  Dobie, Kenneth W.
;  APPLICANT:  Jain, Ravi
;  TITLE OF INVENTION: MODULATION OF CEACAM1 EXPRESSION
;  FILE REFERENCE: ISIS0101-100 (RTS-0655US)
;  CURRENT APPLICATION NUMBER: US/10/889,101
;  CURRENT FILING DATE:  2004-07-12
;  PRIOR APPLICATION NUMBER: 60/486,652
;  PRIOR FILING DATE: 2003-07-12
;  NUMBER OF SEQ ID NOS: 298
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 28920
;   TYPE: DNA
;   ORGANISM: SARS Coronavirus isolate BJ03
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION:
;   OTHER INFORMATION: n = A,T,C or G
US-10-889-101-5

  Query Match            90.9%;  Score 20;  DB 21;  Length 28920;
  Best Local Similarity  75.0%;  Pred. No. 4.6;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        776 GTGAACTCACTCGTGAGCTC 795


RESULT 13
US-10-889-101-6
; Sequence 6, Application US/10889101
; Publication No. US20050107324A1
; GENERAL INFORMATION:
;  APPLICANT: Bennett, C. Frank
;  APPLICANT:  Dobie, Kenneth W.
;  APPLICANT:  Jain, Ravi
;  TITLE OF INVENTION: MODULATION OF CEACAM1 EXPRESSION
;  FILE REFERENCE: ISIS0101-100 (RTS-0655US)
;  CURRENT APPLICATION NUMBER: US/10/889,101
;  CURRENT FILING DATE:  2004-07-12
;  PRIOR APPLICATION NUMBER: 60/486,652
;  PRIOR FILING DATE: 2003-07-12
;  NUMBER OF SEQ ID NOS: 298
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 28920
;   TYPE: DNA
;   ORGANISM: SARS Coronavirus isolate BJ04
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION:
;   OTHER INFORMATION: n = A,T,C or G
US-10-889-101-6

  Query Match            90.9%;  Score 20;  DB 21;  Length 28920;
  Best Local Similarity  75.0%;  Pred. No. 4.6;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        776 GTGAACTCACTCGTGAGCTC 795


RESULT 14
US-10-831-901A-29819
; Sequence 29819, Application US/10831901A
; Publication No. US20050100885A1
; GENERAL INFORMATION:
;  APPLICANT: Crooke, Stanley T.
;  APPLICANT:  Ecker, David J.
;  APPLICANT:  Sampath, Rangarajan
;  APPLICANT:  Freier, Susan M.
;  APPLICANT:  Massire, Christian
;  APPLICANT:  Hofstadler, Steven A.
;  APPLICANT:  Lowery, Kristin Sannes
;  APPLICANT:  Swayze, Eric
;  APPLICANT:  Baker, Brenda F.
;  APPLICANT:  Bennett, C. Frank
;  TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
;  TITLE OF INVENTION:  Acute Respiratory Syndrome (SARS)
;  FILE REFERENCE: ISIS0083-100 (BIOL0008US)
;  CURRENT APPLICATION NUMBER: US/10/831,901A
;  CURRENT FILING DATE:  2004-04-26
;  PRIOR APPLICATION NUMBER: 60/466,426
;  PRIOR FILING DATE: 2003-04-28
;  PRIOR APPLICATION NUMBER: 60/468,562
;  PRIOR FILING DATE: 2003-05-06
;  PRIOR APPLICATION NUMBER: 60/467,770
;  PRIOR FILING DATE: 2003-04-30
;  PRIOR APPLICATION NUMBER: 60/468,627
;  PRIOR FILING DATE: 2003-05-06
;  PRIOR APPLICATION NUMBER: 60/477,637
;  PRIOR FILING DATE: 2003-06-10
;  PRIOR APPLICATION NUMBER: 60/483,579
;  PRIOR FILING DATE: 2003-06-27
;  NUMBER OF SEQ ID NOS: 30063
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29819
;   LENGTH: 29013
;   TYPE: DNA
;   ORGANISM: SARS Coronavirus
US-10-831-901A-29819

  Query Match            90.9%;  Score 20;  DB 21;  Length 29013;
  Best Local Similarity  75.0%;  Pred. No. 4.6;
  Matches   15;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        697 GTGAACTCACTCGTGAGCTC 716


RESULT 15
US-10-831-901A-29742
; Sequence 29742, Application US/10831901A
; Publication No. US20050100885A1
; GENERAL INFORMATION:
;  APPLICANT: Crooke, Stanley T.
;  APPLICANT:  Ecker, David J.
;  APPLICANT:  Sampath, Rangarajan
;  APPLICANT:  Freier, Susan M.
;  APPLICANT:  Massire, Christian
;  APPLICANT:  Hofstadler, Steven A.
;  APPLICANT:  Lowery, Kristin Sannes
;  APPLICANT:  Swayze, Eric
;  APPLICANT:  Baker, Brenda F.
;  APPLICANT:  Bennett, C. Frank
;  TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
;  TITLE OF INVENTION:  Acute Respiratory Syndrome (SARS)
;  FILE REFERENCE: ISIS0083-100 (BIOL0008US)
;  CURRENT APPLICATION NUMBER: US/10/831,901A
;  CURRENT FILING DATE:  2004-04-26
;  PRIOR APPLICATION NUMBER: 60/466,426
;  PRIOR FILING DATE: 2003-04-28
;  PRIOR APPLICATION NUMBER: 60/468,562
;  PRIOR FILING DATE: 2003-05-06
;  PRIOR APPLICATION NUMBER: 60/467,770
;  PRIOR FILING DATE: 2003-04-30
;  PRIOR APPLICATION NUMBER: 60/468,627
;  PRIOR FILING DATE: 2003-05-06
;  PRIOR APPLICATION NUMBER: 60/477,637
;  PRIOR FILING DATE: 2003-06-10
;  PRIOR APPLICATION NUMBER: 60/483,579
;  PRIOR FILING DATE: 2003-06-27
;  NUMBER OF SEQ ID NOS: 30063
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29742
;   LENGTH: 29206
;   TYPE: DNA
;   ORGANISM: SARS Coronavirus
US-10-831-901A-29742

  Query Match            90.9%;  Score 20;  DB 21;  Length 29206;
  Best Local Similarity  75.0%;  Pred. No. 4.6;
```

Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||:||:||||:|
Db        751 GTGAACTCACTCGTGAGCTC 770
```

Search completed: June 24, 2005, 04:35:43
Job time : 259.433 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 23, 2005, 23:49:32 ; Search time 1551.26 Seconds
(without alignments)
539.828 Million cell updates/sec

Title: US-10-848-737-1
Perfect score: 22
Sequence: 1 gugaacucacucgugagcuctt 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	18.8	85.5	254	8	AQ479024	AQ479024 RPCI-11-2
c	2	18.8	85.5	1344	4	BG917388	BG917388 602816903
c	3	18.4	83.6	915	2	BF100955	BF100955 601754406
c	4	18.4	83.6	1580	3	AK031386	AK031386 Mus muscu
	5	17.8	80.9	451	8	AZ312688	AZ312688 1M0028J02
c	6	17.8	80.9	506	5	BP091999	BP091999 BP091999
c	7	17.8	80.9	514	5	BQ820183	BQ820183 1030082F1
c	8	17.8	80.9	514	8	AZ236076	AZ236076 RPCI-23-8
c	9	17.8	80.9	544	1	AV389206	AV389206 AV389206
c	10	17.8	80.9	544	6	CB213780	CB213780 OML04060
c	11	17.8	80.9	546	1	AV619265	AV619265 AV619265
c	12	17.8	80.9	547	8	AZ259445	AZ259445 RPCI-23-1
c	13	17.8	80.9	653	5	BQ824823	BQ824823 1030121E1
c	14	17.8	80.9	658	8	AZ695771	AZ695771 RPCI-23-2
c	15	17.8	80.9	667	5	BQ820182	BQ820182 1030082F1
c	16	17.8	80.9	689	4	BG844655	BG844655 1024007C0
c	17	17.8	80.9	749	9	AG427536	AG427536 Mus muscu
c	18	17.8	80.9	770	2	BF862564	BF862564 963036A08
	19	17.4	79.1	440	8	AQ997054	AQ997054 RPCI-23-3
	20	17.4	79.1	442	8	AQ996162	AQ996162 RPCI-23-3
c	21	17.4	79.1	575	7	CN680938	CN680938 E0144A04-
	22	17.4	79.1	689	9	CE718342	CE718342 tigr-gss-
c	23	17.4	79.1	691	9	CE779822	CE779822 tigr-gss-
	24	17.4	79.1	710	2	BF163602	BF163602 601769987
	25	17.4	79.1	712	5	BU709067	BU709067 UI-M-EW0-
	26	17.4	79.1	732	9	AG401233	AG401233 Mus muscu
	27	17.4	79.1	919	9	CL501597	CL501597 SAIL_700_
c	28	17.2	78.2	246	2	AW466607	AW466607 fi59a01.x
c	29	17.2	78.2	355	7	CR475310	CR475310 CR475310
	30	17.2	78.2	385	1	AA087904	AA087904 mn99h01.r
	31	17.2	78.2	446	9	CL369486	CL369486 RPCI44_29
	32	17.2	78.2	458	1	AL924913	AL924913 AL924913
c	33	17.2	78.2	462	4	BG398975	BG398975 602440832
c	34	17.2	78.2	508	8	AQ382507	AQ382507 RPCI11-16
	35	17.2	78.2	551	5	BQ588777	BQ588777 E012534-0
	36	17.2	78.2	554	9	CL390018	CL390018 RPCI44_28
	37	17.2	78.2	603	6	CB581909	CB581909 AMGNNUC:N
	38	17.2	78.2	636	9	AG074412	AG074412 Pan trogl
	39	17.2	78.2	656	1	AV269391	AV269391 AV269391
	40	17.2	78.2	656	9	CR165590	CR165590 Reverse s
	41	17.2	78.2	703	8	BH595727	BH595727 BOGFZ33TR
	42	17.2	78.2	714	9	BX173618	BX173618 Danio rer
	43	17.2	78.2	749	8	BZ966478	BZ966478 PUDFL16TD
	44	17.2	78.2	758	9	AG423346	AG423346 Mus muscu
c	45	17.2	78.2	813	7	CO960306	CO960306 AGENCOURT

ALIGNMENTS

RESULT 1
AQ479024
LOCUS AQ479024 254 bp DNA linear GSS 23-APR-1999
DEFINITION RPCI-11-254M5.TV RPCI-11 Homo sapiens genomic clone RPCI-11-254M5, genomic survey sequence.
ACCESSION AQ479024
VERSION AQ479024.1 GI:4661143
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 254)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: RPCI-11-254M5.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .254
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7597444"
/db_xref="taxon:9606"
/clone="RPCI-11-254M5"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"
ORIGIN

```
  Query Match             85.5%;  Score 18.8;  DB 8;  Length 254;
  Best Local Similarity   68.2%;  Pred. No. 1.3e+02;
  Matches   15;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:||||:| |:| :||||:|||
Db          1 GTGAACTCTCTCCTGAGCTCTT 22


RESULT 2
BG917388/c
LOCUS       BG917388               1344 bp    mRNA    linear   EST 05-JUN-2001
DEFINITION  602816903F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4945921 5',
            mRNA sequence.
ACCESSION   BG917388
VERSION     BG917388.1  GI:14297864
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1344)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10893  row: f  column: 02
            High quality sequence stop: 309.
FEATURES             Location/Qualifiers
     source          1. .1344
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4945921"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam6"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
ORIGIN

  Query Match             85.5%;  Score 18.8;  DB 4;  Length 1344;
  Best Local Similarity   68.2%;  Pred. No. 1.7e+02;
  Matches   15;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCTT 22
              |:|  |:|||:||:||||:|||
Db        996 GTGGTCTCACTCGTGAGCTCTT 975


RESULT 3
BF100955/c
LOCUS       BF100955               915 bp    mRNA    linear   EST 19-OCT-2000
DEFINITION  601754406F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3983736 5',
            mRNA sequence.
ACCESSION   BF100955
VERSION     BF100955.1  GI:10883481
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9184  row: k  column: 01
            High quality sequence stop: 579.
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3983736"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="3 months, virgin"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam1"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
ORIGIN

  Query Match             83.6%;  Score 18.4;  DB 2;  Length 915;
  Best Local Similarity   70.0%;  Pred. No. 2.6e+02;
  Matches   14;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUC 20
              |:||||:|||: |:||||:|
Db        476 GTGAACTCACTGGTGAGCTC 457


RESULT 4
AK031386/c
LOCUS       AK031386               1580 bp    mRNA    linear   HTC 03-APR-2004
DEFINITION  Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
            enriched library, clone:6030419L17 product:DJ1014D13.2 (NOVEL
            PROTEIN SIMILAR TO ACTN3 (ACTININ, ALPHA 3)) (FRAGMENT) homolog
            [Homo sapiens], full insert sequence.
ACCESSION   AK031386
VERSION     AK031386.1  GI:26327268
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
```

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1580)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.
FEATURES Location/Qualifiers
source 1. .1580
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:6030419L17"
/db_xref="taxon:10090"
/clone="6030419L17"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
CDS 131. .532
/note="unnamed protein product; DJ1014D13.2 (NOVEL PROTEIN SIMILAR TO ACTN3 (ACTININ, ALPHA 3)) (FRAGMENT) homolog [Homo sapiens] (SPTR|Q9UH44, evidence: FASTY, 88.7%ID, 65.2%length, match=372)
putative"
/codon_start=1
/protein_id="BAC27378.1"
/db_xref="GI:26327269"
/translation="MAGPRGALLAWCRRQCEGYRGVDIRDLSSSFRDGLAFCAILHRH RPDLLDFQSLSKENVFENNRLAFEVAEKELGIPALLDPNDMVSMSVPDCLSIMTYVSQ YYNHFTSSGQGERGPSPSGRLPWAVYFFAEQ"
ORIGIN

Query Match 83.6%; Score 18.4; DB 3; Length 1580;
Best Local Similarity 70.0%; Pred. No. 2.8e+02;
Matches 14; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCGUGAGCUC 20
|:||||:|||: |:||||:|
Db 776 GTGAACTCACTGGTGAGCTC 757

RESULT 5
AZ312688
LOCUS AZ312688 451 bp DNA linear GSS 29-SEP-2000
DEFINITION 1M0028J02R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0028J02 R, genomic survey sequence.
ACCESSION AZ312688
VERSION AZ312688.1 GI:10356723
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 451)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0028 row: J column: 02
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 451.
FEATURES Location/Qualifiers
source 1. .451
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0028J02"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to

adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
ORIGIN

Query Match 80.9%; Score 17.8; DB 8; Length 451;
Best Local Similarity 71.4%; Pred. No. 4.7e+02;
Matches 15; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCGUGAGCUCT 21
| ||||:|||: |:||||:||
Db 177 GAGAACTCACTGGTGAGCTCT 197

RESULT 6
BP091999/c
LOCUS BP091999 506 bp mRNA linear EST 30-JUN-2004
DEFINITION BP091999 Chlamydomonas reinhardtii C9 various conditions
Chlamydomonas reinhardtii cDNA clone MX248b10_r 5', mRNA sequence.
ACCESSION BP091999
VERSION BP091999.1 GI:49464086
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 506)
AUTHORS Asamizu,E., Nakamura,Y., Miura,K., Fukuzawa,H., Fujiwara,S.,
Hirono,M., Iwamoto,K., Matsuda,Y., Minagawa,J., Shimogawara,K.,
Takahashi,Y. and Tabata,S.
TITLE Establishment of Publicly Available cDNA Material and Information
Resource of Chlamydomonas reinhardtii (Chlorophyta), to Facilitate
Gene Function Analysis
JOURNAL Phycologia (2004) In press
COMMENT Contact: Erika Asamizu
. The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES Location/Qualifiers
source 1. .506
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="MX248b10_r"
/clone_lib="Chlamydomonas reinhardtii C9 various
conditions"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was made from a mixture of cells
grown under various conditions"
ORIGIN

Query Match 80.9%; Score 17.8; DB 5; Length 506;
Best Local Similarity 66.7%; Pred. No. 4.8e+02;
Matches 14; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 2 UGAACUCACUCGUGAGCUCTT 22
: |||:|| :||:||||:|||
Db 503 TAAACTCATTCGTGAGCTCTT 483

RESULT 7
BQ820183/c
LOCUS BQ820183 514 bp mRNA linear EST 01-AUG-2002
DEFINITION 1030082F11.y2 C. reinhardtii CC-1690, Deflagellation (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BQ820183
VERSION BQ820183.1 GI:22070845
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 514)
AUTHORS Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,
Lefebvre,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1030
JOURNAL Unpublished (2002)
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.
FEATURES Location/Qualifiers
source 1. .514
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Deflagellation
(normalized), Lambda Zap II"
/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; Deflagellation library, constructed by John Davies
and Jeffrey McDermott, combines cDNAs from CC-1690 cells
which had been re-synthesizing flagella for 15, 30 and 60
min after being deflagellated by pH shock. PolyA mRNA was
purified from each sample, pooled and cDNA synthesized.
The cDNA was directionally cloned into lambda Zap II
(Stratagene) in the EcoRI (5') and XhoRI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with ExAssist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al., (1996) Genome Research 6: 791-806."
ORIGIN

Query Match 80.9%; Score 17.8; DB 5; Length 514;
Best Local Similarity 66.7%; Pred. No. 4.8e+02;
Matches 14; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 2 UGAACUCACUCGUGAGCUCTT 22
: |||:|| :||:||||:|||
Db 483 TAAACTCATTCGTGAGCTCTT 463

RESULT 8
AZ236076/c
LOCUS AZ236076 514 bp DNA linear GSS 14-JUN-2000
DEFINITION RPCI-23-84D17.TJ RPCI-23 Mus musculus genomic clone RPCI-23-84D17,
genomic survey sequence.
ACCESSION AZ236076
VERSION AZ236076.1 GI:8544122
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 514)
AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S.,
Akinret,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de
Jong,P. and Fraser,C.M.
TITLE Mouse BAC End Sequences from Library RPCI-23
JOURNAL Unpublished (1999)
COMMENT Other_GSSs: RPCI-23-84D17.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC

```
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 84  row: D  column: 17
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .514
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-84D17"
                     /sex="Female"
                     /lab_host="DH10B"
                     /clone_lib="RPCI-23"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
ORIGIN

  Query Match             80.9%;  Score 17.8;  DB 8;  Length 514;
  Best Local Similarity   71.4%;  Pred. No. 4.8e+02;
  Matches   15;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCT 21
              | ||||:|||: |:||||:||
Db        226 GAGAACTCACTGGTGAGCTCT 206


RESULT 9
AV389206/c
LOCUS       AV389206                544 bp    mRNA    linear   EST 29-SEP-2000
DEFINITION  AV389206 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
            cDNA clone CM040a08_r, mRNA sequence.
ACCESSION   AV389206
VERSION     AV389206.1  GI:6543422
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 544)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S., Fukuzawa,H. and Tabata,S.
  TITLE     A large scale structural analysis of cDNAs in a unicellular green
            alga, Chlamydomonas reinhardtii. I. Generation of 3433
            non-redundant expressed sequence tags
  JOURNAL   DNA Res. 6 (6), 369-373 (1999)
  MEDLINE   20152988
   PUBMED   10691129
COMMENT     Contact: Yasukazu Nakamura
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: ynakamu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .544
                     /organism="Chlamydomonas reinhardtii"
                     /mol_type="mRNA"
                     /strain="C9"
                     /db_xref="taxon:3055"
                     /clone="CM040a08_r"
                     /dev_stage="photoautotrophic growth"
                     /clone_lib="Chlamydomonas reinhardtii C9"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
ORIGIN

  Query Match             80.9%;  Score 17.8;  DB 1;  Length 544;
  Best Local Similarity   66.7%;  Pred. No. 4.8e+02;
  Matches   14;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 UGAACUCACUCGUGAGCUCTT 22
              : |||:|| :||:||||:|||
Db        497 TAAACTCATTCGTGAGCTCTT 477


RESULT 10
CB213780/c
LOCUS       CB213780                544 bp    mRNA    linear   EST 05-FEB-2003
DEFINITION  OML04060 Oryza minuta HybriZAP-2.1 XR library Oryza minuta cDNA 5',
            mRNA sequence.
ACCESSION   CB213780
VERSION     CB213780.1  GI:28259871
KEYWORDS    EST.
SOURCE      Oryza minuta
  ORGANISM  Oryza minuta
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 544)
  AUTHORS   Shin,J.S.
  TITLE     Oryza minuta HybriZAP-2.1 XR library
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Jeong Sheop Shin
            Plant Molecular Genetics
            Graduate School of Biotechnology, University of Korea
            136-701 Anam-dong 5/1 Seoul, Korea
            Tel: 00 82 2 3290 3430
            Fax: 00 82 2 927 9028
            Email: jsshin@kuccnx.korea.ac.kr.
FEATURES             Location/Qualifiers
     source          1. .544
                     /organism="Oryza minuta"
                     /mol_type="mRNA"
                     /db_xref="taxon:63629"
                     /dev_stage="4-weeks after germination"
                     /clone_lib="Oryza minuta HybriZAP-2.1 XR library"
                     /note="Organ: immature leaf"
ORIGIN

  Query Match             80.9%;  Score 17.8;  DB 6;  Length 544;
  Best Local Similarity   66.7%;  Pred. No. 4.8e+02;
  Matches   14;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 UGAACUCACUCGUGAGCUCTT 22
              :|| |:|||:| :||||:|||
Db        199 TGAGCTCACTCTTGAGCTCTT 179


RESULT 11
AV619265/c
LOCUS       AV619265                546 bp    mRNA    linear   EST 15-DEC-2000
DEFINITION  AV619265 Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas
            reinhardtii cDNA clone LC005h05_r 5', mRNA sequence.
ACCESSION   AV619265
VERSION     AV619265.1  GI:10768440
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 546)
  AUTHORS   Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
            Nakamura,Y. and Tabata,S.
  TITLE     Generation of expressed sequence tags from low-CO2 and high-CO2
            adapted cells of Chlamydomonas reinhardtii
  JOURNAL   DNA Res. 7 (5), 305-307 (2000)
  MEDLINE   20539644
   PUBMED   11089912
```

```
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .546
                     /organism="Chlamydomonas reinhardtii"
                     /mol_type="mRNA"
                     /strain="C9"
                     /db_xref="taxon:3055"
                     /clone="LC005h05_r"
                     /clone_lib="Chlamydomonas reinhardtii 5% to 0.04% CO2"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI; The cDNA library was constructed from cells cultured
                     in a carbon stress acclimatized condition in which carbon
                     dioxide concentration in the bubbling gas was changed from
                     5% to 0.04%"
ORIGIN

  Query Match             80.9%;  Score 17.8;  DB 1;  Length 546;
  Best Local Similarity   66.7%;  Pred. No. 4.8e+02;
  Matches   14;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 UGAACUCACUCGUGAGCUCTT 22
              : |||:|| :||:||||:|||
Db        476 TAAACTCATTCGTGAGCTCTT 456


RESULT 12
AZ259445/c
LOCUS       AZ259445                 547 bp    DNA     linear   GSS 26-JUL-2000
DEFINITION  RPCI-23-114N6.TV RPCI-23 Mus musculus genomic clone RPCI-23-114N6,
            genomic survey sequence.
ACCESSION   AZ259445
VERSION     AZ259445.1  GI:9465935
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S.,
            Akinret,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de
            Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-114N6.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 114  row: N  column: 6
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .547
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-114N6"
                     /sex="Female"
                     /lab_host="DH10B"
                     /clone_lib="RPCI-23"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
ORIGIN

  Query Match             80.9%;  Score 17.8;  DB 8;  Length 547;
  Best Local Similarity   71.4%;  Pred. No. 4.8e+02;
  Matches   15;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCT 21
              | ||||:|||: |:||||:||
Db        199 GAGAACTCACTGGTGAGCTCT 179


RESULT 13
BQ824823/c
LOCUS       BQ824823                 653 bp    mRNA    linear   EST 01-AUG-2002
DEFINITION  1030121E11.y1 C. reinhardtii CC-1690, Deflagellation (normalized),
            Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION   BQ824823
VERSION     BQ824823.1  GI:22076047
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 653)
  AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,
            Lefebvre,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
  TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1030
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
FEATURES             Location/Qualifiers
     source          1. .653
                     /organism="Chlamydomonas reinhardtii"
                     /mol_type="mRNA"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690, Deflagellation
                     (normalized), Lambda Zap II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; Deflagellation library, constructed by John Davies
                     and Jeffrey McDermott, combines cDNAs from CC-1690 cells
                     which had been re-synthesizing flagella for 15, 30 and 60
                     min after being deflagellated by pH shock. PolyA mRNA was
                     purified from each sample, pooled and cDNA synthesized.
                     The cDNA was directionally cloned into lambda Zap II
                     (Stratagene) in the EcoRI (5') and XhoRI (3') sites.
                     pBluescript II SK- plasmids were excised from the lambda
                     ZAP clones by superinfection with ExAssist (Stratagene)
                     phage. The library was normalized using method 4 described
                     in Bonaldo et al., (1996) Genome Research 6: 791-806."
ORIGIN

  Query Match             80.9%;  Score 17.8;  DB 5;  Length 653;
  Best Local Similarity   66.7%;  Pred. No. 4.9e+02;
  Matches   14;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 UGAACUCACUCGUGAGCUCTT 22
              : |||:|| :||:||||:|||
```

```
Db        482 TAAACTCATTCGTGAGCTCTT 462


RESULT 14
AZ695771/c
LOCUS       AZ695771                658 bp    DNA     linear   GSS 24-JAN-2001
DEFINITION  RPCI-23-240D2.TV RPCI-23 Mus musculus genomic clone RPCI-23-240D2,
            genomic survey sequence.
ACCESSION   AZ695771
VERSION     AZ695771.1  GI:12410231
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 658)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S.,
            Akinret,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de
            Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-240D2.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 240  row: D  column: 2
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .658
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-240D2"
                     /sex="Female"
                     /lab_host="DH10B"
                     /clone_lib="RPCI-23"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
ORIGIN

  Query Match            80.9%;  Score 17.8;  DB 8;  Length 658;
  Best Local Similarity  71.4%;  Pred. No. 4.9e+02;
  Matches   15;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GUGAACUCACUCGUGAGCUCT 21
              | ||||:|||: |:||||:||
Db        228 GAGAACTCACTGGTGAGCTCT 208


RESULT 15
BQ820182/c
LOCUS       BQ820182                667 bp    mRNA    linear   EST 01-AUG-2002
DEFINITION  1030082F11.y1 C. reinhardtii CC-1690, Deflagellation (normalized),
            Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION   BQ820182
VERSION     BQ820182.1  GI:22070844
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 667)
  AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,
            Lefebvre,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
  TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1030
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
FEATURES             Location/Qualifiers
     source          1. .667
                     /organism="Chlamydomonas reinhardtii"
                     /mol_type="mRNA"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690, Deflagellation
                     (normalized), Lambda Zap II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; Deflagellation library, constructed by John Davies
                     and Jeffrey McDermott, combines cDNAs from CC-1690 cells
                     which had been re-synthesizing flagella for 15, 30 and 60
                     min after being deflagellated by pH shock. PolyA mRNA was
                     purified from each sample, pooled and cDNA synthesized.
                     The cDNA was directionally cloned into lambda Zap II
                     (Stratagene) in the EcoRI (5') and XhoRI (3') sites.
                     pBluescript II SK- plasmids were excised from the lambda
                     ZAP clones by superinfection with ExAssist (Stratagene)
                     phage. The library was normalized using method 4 described
                     in Bonaldo et al., (1996) Genome Research 6: 791-806."
ORIGIN

  Query Match            80.9%;  Score 17.8;  DB 5;  Length 667;
  Best Local Similarity  66.7%;  Pred. No. 5e+02;
  Matches   14;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 UGAACUCACUCGUGAGCUCTT 22
              : |||:|| :||:||||:|||
Db        483 TAAACTCATTCGTGAGCTCTT 463


Search completed: June 24, 2005, 04:05:34
Job time : 1564.43 secs
```

THIS PAGE IS BLANK